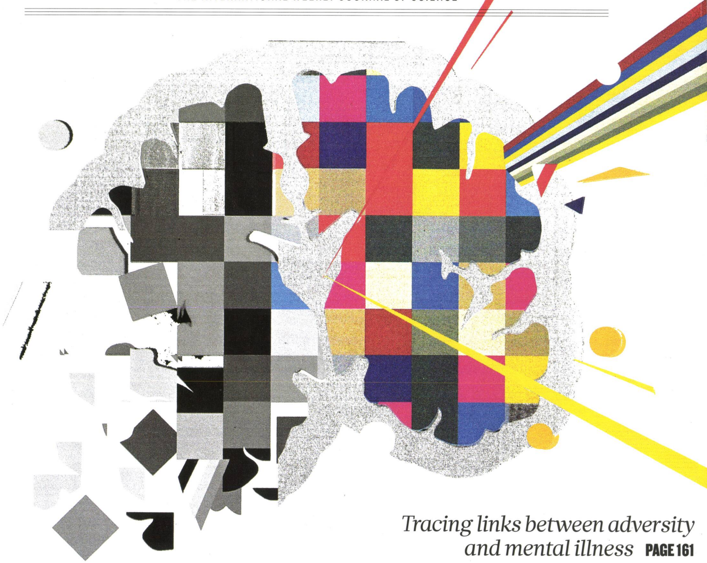
OUTLOOK
Physics masterclass

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



Tracing links between adversity and mental illness **PAGE 161**

STRESS AND RESILIENCE

HEALTH

SLEEPERS ARISE

Why notions of 'normal' sleep need a rethink

PAGE 173

TECHNOLOGY

GRAPHENE'S TIME HAS COME

Bulk production paves the way for new applications

PAGE 192

PALAEONTOLOGY

COMPOUND INTEREST

Eye and brain complexity in a Cambrian fossil

PAGE 258

➲ NATURE.COM/NATURE

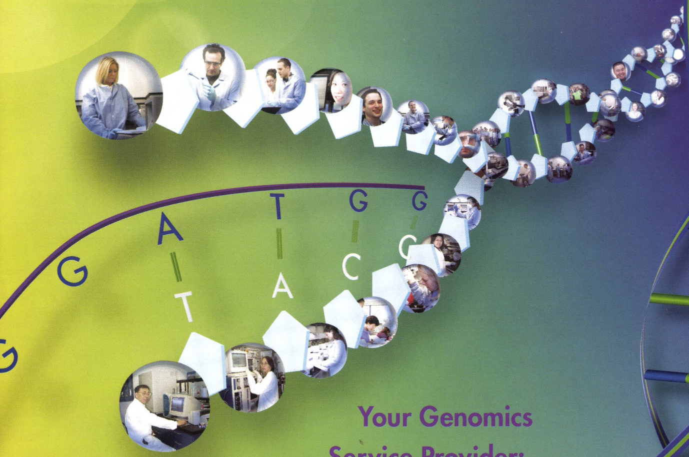
11 October 2012 £10

Vol. 490, No. 7419

GENEWIZ®
Solid science. Superior service.

Contributing to the Advancement of Global Life Sciences

Specializing in Genomics Since 1999



Your Genomics Service Provider:

- Market Leader in Sanger DNA Sequencing
- High Quality Gene Synthesis
- Next Generation Sequencing
- Translational Genomics
- CLIA Certified & GxP Compliant

www.genewiz.com

EPITOMICS®
an abcam® company

Now Available
4000 RABBIT MONOCLONALS
High Quality Antibodies

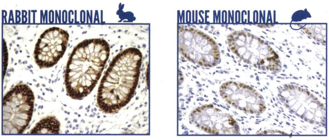
MONO

THE RABBIT MONOCLONAL ADVANTAGE

- More diverse epitope recognition
- High specificity and affinity
- Improved response to mouse antigens
- Highly validated - every antibody tested in WB, IHC, ICC, Flow Cytometry, and IP

Rabbit Monoclonal Antibodies (RabMAbs®) provide the combined benefits of superior antigen recognition of the rabbit immune system with the specificity and consistency of a monoclonal antibody, bringing you the highest quality antibody possible.

Over ***4000 RabMAbs***® online, visit www.epitomics.com



RABBIT MONOCLONAL
MOUSE MONOCLONAL

CDX2 RabMAb
CDX2 Mouse MAb

Paraffin-embedded human colon tissue stained with Epitomics' CDX2 RabMAb (Cat.# 2475-1) and Vendor A's CDX2 Mouse Monoclonal under optimal conditions

Find out more @ www.epitomics.com/advantage

100% SATISFACTION GUARANTEE *Antibody Satisfaction Guarantee*

www.epitomics.com

US & Canada | 1-877-772-2622
Outside NA | 650-583-6688

GOT AN IDEA BREWING?
MAKE IT HAPPEN WITH A
CATALYST GRANT OF UP TO $25k

If you've got an idea
that blends science with
technology, we've got the
funding and resources
to make it happen.

Get in touch at **digital-science.com/catalyst**

DIGITAL
science

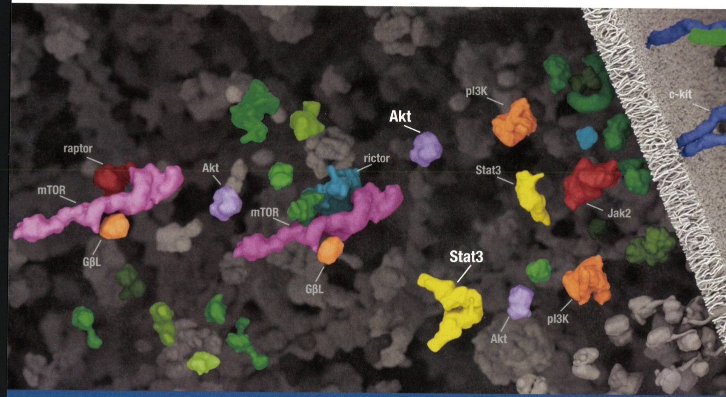
F·S·T®
FINE SCIENCE TOOLS

FINE SURGICAL INSTRUMENTS FOR RESEARCH™

stainless germany
F·S·T 14094-11
F·S·T
11..09

Incredible discoveries with Fine Science Tools

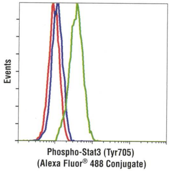
SHIPPING GLOBALLY SINCE 1974
REQUEST A CATALOG AT FINESCIENCE.COM OR CALL 1-800-521-2109



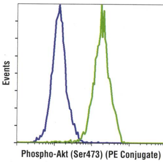
Conjugated Antibodies for Flow Cytometry *from Cell Signaling Technology*

Unparalleled product quality, validation, and technical support

- :: The highest quality antibodies provide you with the brightest signal and the lowest background.
- :: Extensive validation by our in-house flow cytometry group means optimization is not left up to you, the user.
- :: Technical support provided by our flow cytometry specialists translates into a thorough, fast, and accurate response.
- :: High quality custom conjugations of our off-the-shelf antibodies are also available, please inquire about sizing and pricing options.
- :: Multiple fluorochrome options are available to meet your unique multiplexing needs.



Phospho-Stat3 (Tyr705) (D3A7) XP® Rabbit mAb (Alexa Fluor® 488 Conjugate) #4323: Flow cytometric analysis of Jurkat cells, untreated (blue) or IFN-α treated (green), using #4323 compared to nonspecific control antibody (red).



Phospho-Akt (Ser473) (D9E) XP® Rabbit
(PE Conjugate) #5315: Flow cytometric an
of Jurkat cells, untreated (green) or treated w
LY294002 (PI3 Kinase Inhibitor) #9901, Wor
(PI3 Kinase Inhibitor) #9551, and U0126 (ME
Inhibitor) #9903 (blue), using #5315.

for quality products you can trust...

www.cellsignal.com

Cell Signaling

CONTENTS

11 October 2012 / Vol 490 / Issue No 7419

STRESS AND RESILIENCE

EDITORIAL

143 **Life stresses**
Sociologists and biologists need to join forces to study the effects of stress

FEATURES

161 **Under pressure**
A *Nature* special looks at the links between stress and mental illness

162 **Urban decay**
The stress of city life could be increasing the rate of psychosis

165 **The roots of resilience**
Why some people can recover from traumatic events and others can't

COMMENT

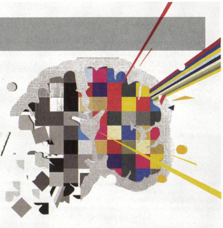
169 **Too toxic to ignore**
Elizabeth H Blackburn & Elissa S Epel
Shortened telomeres could be indicator of stress and disease

171 **Stress makes its molecular mark**
Eric J Nestler
Epigenetics could be key to the differing responses to trauma

BOOKS & ARTS

173 **Broken dreams**
Meredith Wadman

174 **The dispossessed**
Amy Maxmen



COVER ILLUSTRATION BY PADDY MILLS

CAREERS

299 **Under a cloud**
Depression is rife among university students — counselling can help

THIS WEEK

EDITORIALS

143 GENOMICS
Share alike
Data sharing is essential to progress

144 GENOECONOMICS
Fighting chance
Geneticists and economists need to take each other more seriously

WORLD VIEW

145 **Expert assessments can work if lessons are learned**
Ehsan Masood
The Oxford Martin Commission must take heed of others' mistakes

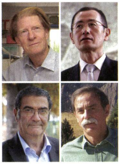
RESEARCH HIGHLIGHTS

146 SELECTIONS FROM THE SCIENTIFIC LITERATURE
House spider designs web for type of prey / Fertile mouse eggs derived from stem cells / Slime moulds leave sticky trails to mark their way / Earth's continents are drifting westwards

SEVEN DAYS

148 THE NEWS IN BRIEF
Mammoth remains found in Siberia / Hope for muscular dystrophy drug / Radio-telescope opens in Australia / Nuclear safety concerns in Europe

NEWS IN FOCUS



151 MEDICINE
Nobel prize awarded for cell reprogramming

152 PHYSICS
Quantum optics takes top award

153 FORENSICS
Boston crime-lab scandal exposes system overload

154 INTERDISCIPLINARY RESEARCH
Economics paper that uses population-genetics data sparks row with geneticists

157 BUSINESS
Genome interpreter close to market

158 BIOTECHNOLOGY
Regulators assail study linking GM maize and cancer

159 FUNDING
National Science Foundation policy change rankles some life scientists

TECHNOLOGY

293 HIGH-THROUGHPUT ANATOMY
Charting the brain's networks
The field of connectomics is helping neuroscientists to generate huge amounts of data

COMMENT

BOOKS & ARTS

175 BOOKS IN BRIEF

CORRESPONDENCE

176 Elite grant recipients have nothing to prove / Visa boosts Ireland / Lundmark saw Universe expanding in 1924 / Australian ecosystems threatened

FUTURES

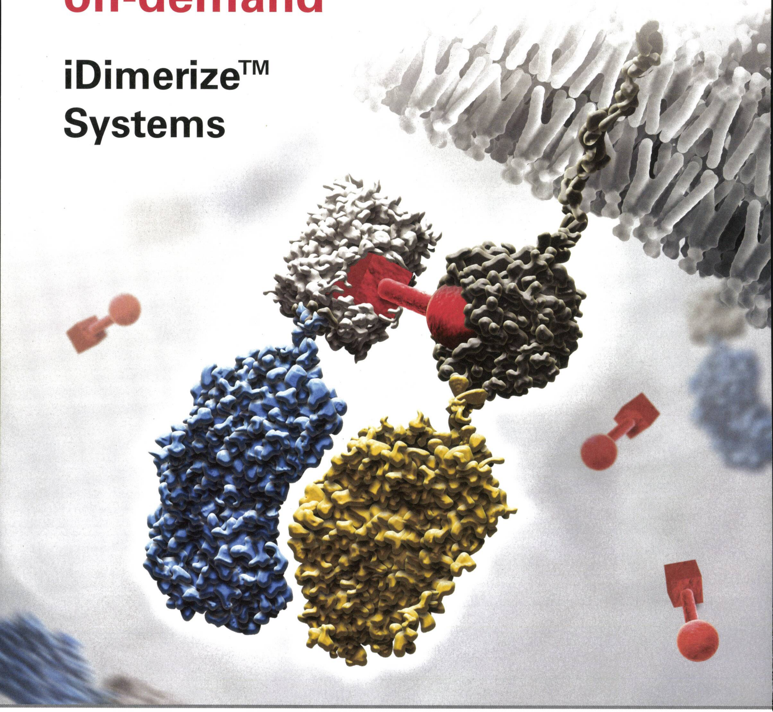
304 **Temporal ventures robbed me**
Scott C Mikula

CAREERS

301 COLUMN
Advising on the edge
Katherine Mackey
Interdisciplinary graduate programmes create additional pressure for mentors

NATUREJOBS ADVERTISING FEATURE

Spotlight on neuroscience



Induce protein interactions on-demand

iDimerize™ Systems

Visit www.clontech.com/dimerize
or call 1.800.662.2566

Clontech

Clontech Laboratories, Inc.
A Takara Bio Company

United States/Canada: +1.800.662.2566
For Research Use Only. Not for use in diagnostic or therapeutic procedures. Not for resale. Clontech, the Clontech logo, and iDimerize are trademarks of Clontech Laboratories, Inc. All other marks are the property of their respective owners. Certain trademarks may not be registered in all jurisdictions.
©2012 Clontech Laboratories, Inc:

9.12 US (633521) www.clontech.com

CONTENTS

11 October 2012 / Vol 490 / Issue No 7419

RESEARCH

NEW ONLINE

177 Papers published this week at nature.com

NEWS & VIEWS

178 **NEUROSCIENCE**
Promiscuous vesicles
Dopaminergic neurons can co-release dopamine and GABA
John T Williams **SEE LETTER P.262**

179 **ORGANIC CHEMISTRY**
Single molecules put a ring on it
A cycloaddition reaction for making arynes
John T S Yeoman & Sarah E Reisman
SEE ARTICLE P.208

180 **PALAEONTOLOGY**
Cambrian nervous wrecks
Preserved nervous-system tissue shows *Fuxianhuia* had a tripartite brain
Graham E Budd **SEE LETTER P.258**

182 **ANIMAL MIGRATION**
Catching the wave
Migrating red deer tend to jump ahead of seasonal green-up
John M Fryxell & Tal Avgar

183 **EARTH SCIENCE**
When an oceanic tectonic plate cracks
An analysis of the 11 April 2012 earthquakes
Jean-Yves Royer **SEE LETTERS P.240, P.245 & P.250**

185 **BRAIN DEVELOPMENT**
The neuron family tree remodelled
Different types of neuronal progenitor cells exist in the neocortex
Oscar Marín

PERSPECTIVE

187 **METHODS A call for transparent reporting to optimize the predictive value of preclinical research**
S C Landis et al.

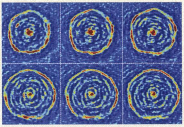
REVIEWS

192 **TECHNOLOGY A roadmap for graphene**
K S Novoselov et al.

201 **NEUROSCIENCE Molecular machines governing exocytosis of synaptic vesicles**
R Jahn & D Fasshauer

ARTICLES

208 **CHEMISTRY The hexadehydro-Diels–Alder reaction**
T R Hoye et al. **SEE N&V P. 179**



ASTRONOMY

Missed mass

A closer look at R Sculptoris, a star at a late stage of stellar evolution **PAGE 232**

213 **CELL BIOLOGY Small heat-shock proteins protect from heat-stroke-associated neurodegeneration**
N Kourtis, V Nikoletopoulou & N Tavernarakis

219 **NEUROSCIENCE Retinal waves coordinate patterned activity throughout the developing visual system**
J B Ackman, T J Burbridge & M C Crair

226 **NEUROSCIENCE A neural circuit for spatial summation in visual cortex**
H Adesnik et al.

LETTERS

232 **ASTRONOMY Unexpectedly large mass loss during the thermal pulse cycle of the red giant star R Sculptoris**
M Maercker et al.

235 **MATERIALS SCIENCE Probing graphene grain boundaries with optical microscopy**
D L Duong et al.

240 **EARTH SCIENCE April 2012 intra-oceanic seismicity off Sumatra boosted by the Banda-Aceh megathrust**
M Delescluse et al. **SEE N&V P.183**

245 **EARTH SCIENCE *En échelon* and orthogonal fault ruptures of the 11 April 2012 great intraplate earthquakes**
H Yue, T Lay & K D Koper **SEE N&V P.183**

250 **EARTH SCIENCE The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide**
F F Pollitz, R S Stein, V Sevilgen & R Bürgmann **SEE N&V P.183**

254 **ENVIRONMENTAL SCIENCE Closing yield gaps through nutrient and water management**
N D Mueller et al.

258 **PALAEONTOLOGY Complex brain and optic lobes in an early Cambrian arthropod**
X Ma, X Hou, G D Edgecombe & N J Strausfeld **SEE N&V P.180**

262 **NEUROSCIENCE Dopaminergic neurons inhibit striatal output through non-canonical release of GABA**
N X Tritsch, J B Ding & B L Sabatini
SEE N&V P.178

267 **GENETICS *FTO* genotype is associated with phenotypic variability of body mass index**
J Yang et al.

273 **ANIMAL BEHAVIOUR Controlling interneuron activity in *Caenorhabditis elegans* to evoke chemotactic behaviour**
A Kocabas, C-H Shen, Z V Guo & S Ramanathan

278 **STEM CELLS Restoration of auditory evoked responses by human ES-cell-derived otic progenitors**
W Chen et al.

283 **IMMUNOLOGY HIV-infected T cells are migratory vehicles for viral dissemination**
T T Murooka et al.

288 **IMMUNOLOGY Caspase-11 increases susceptibility to *Salmonella* infection in the absence of caspase-1**
P Broz et al.

292 **CORRIGENDUM Past extreme warming events linked to massive carbon release from thawing permafrost**
R M DeConto et al.

292 **ERRATUM Integrated genomic analyses of ovarian carcinoma**
The Cancer Genome Atlas Research Network

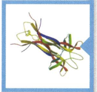
292 **CORRIGENDUM Mitochondrial DNA that escapes from autophagy causes inflammation and heart failure**
T Oka et al.

Neuro Biomarkers

Search

CLIA CERTIFIED

SUPPORTS
GLP STUDIES



Brain-Derived Neurotrophic Factor

✓Friends Message ▾

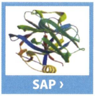
Hometown: Endoplasmic Reticulum

Likes: Keeping neurons happy, cycling, cigars

Dislikes: Depression, air-pollution

Hobbies: Neurogenesis

Looking for: Some action in the hippocampus



NCAM

✓Friends Message ▾

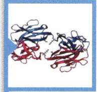
Hometown: Neuroville

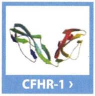
Likes: Learning new things, Brussels sprouts

Dislikes: Breaking up

Hobbies: Cellular matchmaking

Looking for: Long-term potentiation and a neuron to grow old with





Send us your samples.

Find the biomarker profiles that matter.

MYRIAD RBM®

The quantitative multiplexed immunoassay service lab

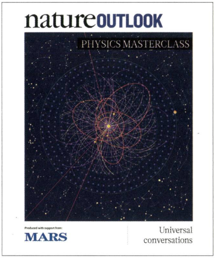
Find out more at myriadrbm.com/biomarker-profiles/Nature

FIND THE
BIOMARKERS YOUR
MOTHER WARNED
YOU ABOUT HERE

OUTLOOK

Supplement to 11 October 2012 / Vol 490 / Issue No. 7419

PHYSICS MASTERCLASS



Turn to page 176

S2 COSMOLOGY
Out of the darkness
Dark energy is the force thought to be driving the Universe outwards, but does it actually exist?

S5 ASTRONOMY
A conversation about observation
Two astronomers, one celebrated the other just starting out, discuss the life of a star gazer

S8 Q&A
Finding fascination in nature
Douglas Osheroff explains how playing with electricity as a boy led to ultralow temperature physics

We talk with five of the laureates at the 2012 Lindau Nobel Laureates Meeting in Germany (page S8).

S10 Q&A
Coming to terms with the Higgs
Martinus Veltman talks about life after Higgs and the loneliness of a theoretical physicist

S11 Q&A
In search of ignorance
William Phillips shares his excitement for beautiful experiments in physics and explains the value of knowing little

S13 Q&A
From Los Alamos to Stockholm
He built his first telescope at age 12, Roy Glauber tells of his stellar career and why Nobels shouldn't be a motivation

S14 Q&A
Socially responsible science
Mario Molina has advice for young scientists on dealing with climate change deniers

MORE ONLINE

NATURE.COM

- *Nature Video* presents a series of conversations caught on camera at this year's Lindau Meeting of Nobel Laureates: **go.nature.com/hm6azq**

COLLECTION

- Visit our online edition, which includes must read articles for anybody considering a career in physics: **go.nature.com/zdjbLp**

nature

11 October 2012 / Vol 490 / Issue No. 7419

EDITORIAL

LONDON, MUNICH, WASHINGTON DC, NEW YORK, BOSTON, PARIS, SAN FRANCISCO, TOKYO nature@nature.com

EDITOR-IN-CHIEF: Philip Campbell
EXECUTIVE EDITOR: Nick Campbell
PUBLISHING EXECUTIVE EDITOR: Maxine Clarke **EXECUTIVE EDITOR:** Véronique Kiermer
CHIEF MAGAZINE EDITOR: Tim Appenzeller
THIS WEEK/NEWS/FEATURES: Alison Abbott, David Adam, Ananyo Bhattacharya, Geoff Brumfiel, Declan Butler, Ewen Callaway, Erika Check Hayden, Daniel Cressey, David Cyranoski, Natasha Gilbert, Eric Hand, Heidi Ledford, Corie Lok, Brendan Maher, Emma Marris, Richard Monastersky, Brian Owens, Helen Pearson, Mark Peplow, Eugenie Samuel Reich, Ivan Semeniuk, Quirin Schiermeier, Jeff Tollefson, Richard Van Noorden, Meredith Wadman, M Mitchell Waldrop **COMMENT:** Sara Abdulla, Joanne Baker, Rosalind Cotter, Barbara Kiser, Alison McCook, Lucy Odling-Smee, Sarah Tomlin **NATURE PODCAST/VIDEO:** Adam Rutherford, Thea Cunningham, Kerri Smith, Charlotte Stoddart **NEWS AND VIEWS:** Sadaf Shadan, Ana Lopes, Andrew Mitchinson, Cesar Sanchez, Marian Turner **RESEARCH:** Ritu Dhand, Karl Ziemelis, Francesca Cesari, I-han Chou, Tanguy Chouard, Rosamund Daw, Alex Eccleston, Angela Eggleston, Joshua Finkelstein, Henry Gee, Patrick Goymer, Noah Gray, Marie-Thérèse Heemels, Magdalena Helmer, Karen Howell, Claudia Lupp, Barbara Marte, Andrew Mitchinson, Juliane Mössinger, Deepa Nath, Wayne Peng, Leslie Sage, Magdalena Skipper, Clare Thomas, Maria Trajkovska, John VanDecar, Liesbeth Venema, Ursula Weiss, Michael White **INSIGHTS/REVIEWS/PERSPECTIVES:** Ursula Weiss **CAREERS:** Gene Russo, Karen Kaplan **TECHNOLOGY FEATURES:** Monya Baker **FUTURES:** Colin Sullivan **SUPPLEMENTS:** Herb Brody, Michelle Grayson, Tony Scully **SUBEDITORS:** Colin Sullivan, Mary Abraham, Sarah Archibald, Nicola Bailey, Nicolas Fanget, Isobel Flanagan, Paul Fletcher, Anne Haggart, Sasha Lewis, Dinah Loon, Louisa Lyon, Anna Novitzky, Yukie Ozawa, Katherine Perry, David Price, Jenny Rooke, David Winters, Anna York **PRODUCTION:** Jenny Henderson, Leonora Dawson-Bowling, Dan Foley, Stewart Fraser, Susan Gray, Simon Gribbin, Hazel Mayhew, James McQuat, Emilia Orviss, Yvonne Strong, Robert Sullivan, Roberta Tarran, Charles Wenz **ART AND DESIGN:** Kelly Buckheit Krause, Wesley Fernandes, Madeline Hutchinson, Barbara Izdebska, Paul Jackman, Jasiek Krzysztofiak, Annthea Lewis, Chris Ryan, Nik Spencer, Claire Welsh, Mark Wood **JAPAN EDITION:** Eiji Matsuda, Mika Ishida, Mitsuyo Utsugi, Sayaka Ayame **PRESS OFFICE:** Ruth Francis, Neda Afsarmanesh, Lisa Boucher, Rachel Twinn, Rebecca Walton **ADMINISTRATION:** Paul Sng, Liz Bankes, Roseann Campbell, Peter Cary, Ffion Cleverley, Dana Coles, Faye Fornasier, Ashlee Gagui, Peter Hewitt, Caroline McLean, Anastasia Panoutsou, Kenneth Simpson, Michelle Viscuso, Diane Yorke

PUBLISHING

LONDON feedback@nature.com

MANAGING DIRECTOR & PUBLISHER: Steven Inchcoombe
ASSOCIATE PUBLISHER: Arash Hejazi
PUBLISHING PROJECT MANAGERS: Claudia Deasy, Christian Manco

TOKYO feedback@natureasia.com

REGIONAL MANAGING DIRECTOR — ASIA-PACIFIC: David Swinbanks
DIRECTOR — ASIA-PACIFIC: Antoine E Bocquet
OPERATIONS DIRECTOR — ASIA-PACIFIC: Hiroshi Minemura

DISPLAY ADVERTISING, SPONSORSHIP & NATUREJOBS

GLOBAL HEAD: Andrew Douglas a.douglas@nature.com

DISPLAY ADVERTISING AND SPONSORSHIP

GLOBAL HEAD: Gerard Preston g.preston@nature.com
ASSISTANT SPONSORSHIP MANAGER: Reya Silao Tel: +44 (0)20 7843 4977 (r.silao@nature.com)

EUROPE/REST OF WORLD display@nature.com Tel: +44 (0)20 7843 4960

GERMANY/SWITZERLAND/AUSTRIA: Sabine Hugi-Fürst Tel: +41 52761 3386
UK/IRELAND/SCANDINAVIA/SPAIN/PORTUGAL: Evelina Rubio Hakansson Tel: +44 (0)20 7843 4079
FRANCE/BELGIUM/THE NETHERLANDS/LUXEMBOURG/ITALY/OTHER EUROPE: Nicola Wright Tel: +44 (0)20 7843 4959

NORTH AMERICA display@natureny.com

NEW ENGLAND/ EASTERN CANADA: Sheila Reardon Tel: +1 617 494 4900
NEW YORK/MID-ATLANTIC/SOUTHEAST: Jim Breault Tel: +1 212 726 9334
MIDWEST: Mike Rossi Tel: +1 212 726 9255
WEST COAST/ WESTERN CANADA: George Lui Tel: +1 415 781 3804
WEST COAST NORTH: Kayla McCutchan Tel: +1 415 403 9038

ASIA-PACIFIC display@natureasia.com

ASIA-PACIFIC HEAD: Kate Yoneyama
JAPAN: Ken Mikami Tel: +81 3 3267 8765
GREATER CHINA/SINGAPORE: Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 124 288 1057

SPONSORSHIP

BUSINESS DEVELOPMENT MANAGER: David Bagshaw Tel: +1 212 726 9215
BUSINESS DEVELOPMENT MANAGER: Patrick Murphy Tel: +1 617 475 9216

NATUREJOBS

naturejobs@nature.com

EUROPEAN HEAD OFFICE, LONDON: Tel: +44 (0)20 7843 4961
EUROPEAN SALES MANAGER: Nils Moeller (n.moeller@nature.com) **BUSINESS DEVELOPMENT MANAGER, MIDDLE EAST:** Jon Giuliani (j.giuliani@nature.com) **ADVERTISING PRODUCTION MANAGER:** Stephen Russell **NATUREJOBS ONLINE PRODUCTION:** Dennis Chu
US HEAD OFFICE, NEW YORK: naturejobs@natureny.com **US SALES MANAGER:** Ken Finnegan
JAPAN HEAD OFFICE, TOKYO: Tel: +81 3 3267 8765 **ASIA-PACIFIC HEAD:** Kate Yoneyama
ASIA-PACIFIC SALES & BUSINESS DEVELOPMENT MANAGER: Yuki Fujiwara **BUSINESS DEVELOPMENT MANAGER, GREATER CHINA, SINGAPORE:** Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 1242881057 (v.chawla@nature.com)

MARKETING & SUBSCRIPTIONS

USA/CANADA/LATIN AMERICA subscriptions@us.nature.com

Tel: (USA/Canada)+1 866 363 7860; (outside USA/Canada) +1 212 726 9223
MARKETING: Sara Girard **FULFILMENT:** Karen Dawson

JAPAN/CHINA/KOREA subscriptions@natureasia.com

Tel: +81 3 3267 8751 **MARKETING:** Sachiyo Ikeda **FULFILMENT:** Masayo Tanaka

EUROPE/REST OF WORLD subscriptions@nature.com

Tel: +44 (0)1256 329242
MARKETING: Elena Woodstock, Marcus Parker, Hannah Phipps **FULFILMENT:** Jean Burrows

INDIA npgindia@nature.com

Tel: +91 124 2881053/54
HEAD OF BUSINESS DEVELOPMENT, INDIA: Debashish Brahmachari **MARKETING:** Harpal Singh Gill

ANNUAL SUBSCRIPTIONS (UK/Europe, including post and packing)

FULL RATE: £1,400
PERSONAL RATES: UK £135; Europe €209; Rest of World £210
STUDENT RATES: UK £79; Europe €122; Rest of World £125
Individual rates apply to payments by personal cheque or credit card. Back issues: £20.00.

ANNUAL SUBSCRIPTIONS (USA, including post and packing)

INSTITUTIONAL/CORPORATE RATE: $2,730 ; **PERSONAL RATE:** $199; **STUDENT RATE:** $99; **POSTDOC RATE:** $119

Printed in USA. Individual rates available only to subscribers paying by personal check or credit card. Orders for student/postdoc subscriptions must be accompanied by a copy of student ID. Rates apply to USA, Canada, Mexico/Central & South America. Add 7% GST tax in Canada (Canadian GST number 140911595). **BACK ISSUES:** US$20.00.

ANNUAL SUBSCRIPTIONS (Asia, including post and packing)

JAPAN: **年間購読料金：** *Nature* のみ 437,800 円 +21,890 円（消費税）=459,690 円（合計）
Nature + *Nature Digest* 466,550 円 +23,328 円（消費税）=489,878（合計）
お問い合わせ：日本出版貿易株式会社 国内営業 1 部
〒101-0064 東京都千代田区猿楽町1-2-1 Tel: 03-3292-3767 Fax: 03-3292-8766
個人/学生購読について
NPG ネイチャー アジア・パシフィック　フルフィルメント & カスタマーサービス
〒162-0843 東京都新宿区市谷田町2-37 千代田ビル Tel: 03-3267-8759 Fax: 03-3267-8752
CHINA: www.natureasia.com/ch/subscribe; **KOREA:** www.natureasia.com/korea/subscribe

NATURE ONLINE

WWW.NATURE.COM/NATURE

CHIEF TECHNOLOGY OFFICER: Howard Ratner **ASSOCIATE DIRECTOR, NATURE.COM:** Daniel Pollock **PLATFORM DEVELOPMENT:** Jeremy Macdonald **WEB PRODUCTION:** Alexander Thurrell, Deborah Anthony **WEB DESIGN:** Juan-Carlos Sobrino, Luke Stavenhagen **PLATFORM INTERFACES:** Amanda Ward

SITE LICENCES, FULFILMENT & CUSTOMER SERVICES

SITE LICENCES: npg.nature.com/libraries feedback@nature.com
CHIEF OPERATIONS OFFICER: John Carroll **ASSOCIATE DIRECTOR OF OPERATIONS:** Dominic Pettit
CUSTOMER SERVICE: Liz Davidson **ORDER PROCESSING OPERATIONS:** Steve Poulter

NATURE PUBLISHING GROUP OFFICES

LONDON: The Macmillan Building, 4 Crinan Street, London N1 9XW Tel: +44 (0)20 7833 4000 Fax: +44 (0)20 7843 4596/7
BASINGSTOKE: Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants RG21 6XS, UK Tel: +44 (0)1256 329242
WASHINGTON DC: 968 National Press Building, 529 14th St NW, Washington DC 20045-1938 Tel: +1 202 737 2355
NEW YORK: 75 Varick St, 9th Floor, New York, NY 10013-1917 Tel: +1 212 726 9200
BOSTON: 25 First Street, Suite 104, Cambridge, MA 02141 Tel: +1 617 475 9275
SAN FRANCISCO: 225 Bush Street, Suite 1453, San Francisco, CA 94104 Tel: +1 415 403 9027
MUNICH: Nymphenburger Strasse 14, D-80335 München Tel: +49 89 549057-13
PARIS: Moulin de Malicorne, Sainte Maure de Touraine, Paris 37800 Tel: +33 2 47 35 72 15
TOKYO: Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843 Tel: +81 3 3267 8751
INDIA: Nature Publishing Group, 3A, 4th Floor, DLF Corporate Park, Gurgaon 122002 Tel: +91 124 2881053/54

Nature® (ISSN 0028-0836) is published by Nature Publishing Group, a division of Macmillan Publishers Ltd (The Macmillan Building, 4 Crinan Street, London N19XW). Registered as a newspaper at the British Post Office. North and South American orders to subscriptions@natureny.com (Tel: +1 866 363 7860). Other orders to subscriptions@nature.com (Tel. +44 (0)1256 329242; Fax +44 (0)1256 812358). Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the relevant copyright fee is paid direct to CCC, 222 Rosewood Drive, Danvers MA 01923, USA. Identification code for *Nature*: 0028-0836/03. CPC PUB AGREEMENT #40032744. In the US, *Nature* (ISSN 0028-0836) is published weekly on Thursday, except the last week in December by Nature Publishing Group, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. US Periodicals postage paid at New York NY, and additional mailing post offices. US POSTMASTER: send address changes to *Nature*, Subscription Dept, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. Published in Japan by NPG Nature Asia-Pacific, Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan. © 2012 Macmillan Publishers Limited. All rights reserved.

eppendorf

EPPENDORF AWARD – FOR YOUNG EUROPEAN INVESTIGATORS –

Winner 2012
Elizabeth Murchison,
United Kingdom

APPLICATION DEADLINE
15 JANUARY 2013

Call for Entries!

Eppendorf Award 2013 for Young European Investigators
Since 1995 the Eppendorf Young Investigator Award is granted annually to European researchers not older than 35 years. It acknowledges outstanding contributions to biomedical research in Europe based on methods of molecular biology, including novel analytical concepts. The winner is selected by an independent expert committee chaired by Reinhard Jahn (Max Planck Institute for Biophysical Chemistry, Göttingen, Germany).

Selection criteria
1. Ability to work independently
2. Creativity
3. Successful practical application of research approach
4. Significance of research results for the field of medicine
5. Achievement of candidate in relation to his/her age

You could be next to receive
> a prize money of 15,000 Euro
> an invitation to the prize ceremony at the EMBL Advanced Training Centre in Heidelberg on 6 June 2013
> an invitation to visit Eppendorf AG in Hamburg
> coverage of your work by Nature in print and in a podcast

Apply online at www.eppendorf.com/award/application

Eppendorf® is a registered trademark of Eppendorf AG. All rights reserved, incl. graphics and images. Copyright ©2012 by Eppendorf AG.

The Eppendorf Award is presented in partnership with
nature

NEW ENGLAND
BioLabs® Inc.
enabling technologies in the life sciences

Stick together.

New DNA Ligases and Ligase Master Mixes

New England Biolabs offers the most extensive selection of high-quality and performance-optimized DNA ligases and ligase master mixes to streamline your cloning experiments.

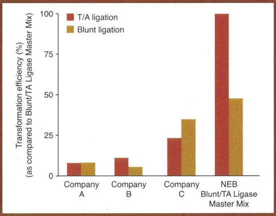
Our expanded portfolio now includes:

- Blunt/TA Ligase Master Mix, optimized for blunt-end and single-base overhang substrates
- Instant Sticky-end Ligase Master Mix, uniquely formulated for the rapid ligation of sticky-end substrates
- T7 DNA Ligase, specific for sticky ends
- ElectroLigase,™ directly compatible with electroporation

Stick together with DNA Ligases and Ligase Master Mixes from NEB.

To request a FREE SAMPLE of our new DNA Ligase Master Mixes, visit **NEBStickTogether.com**

Blunt/TA Ligase Master Mix outperforms the competition



Duplicate ligation reactions of blunt or T/A vector/insert pairs were set up according to the master mix vendors' suggestions. Equal amounts of ligated DNA were used to transform NEB 10-beta Competent E. coli (NEB #C3019) and triplicate plating was performed. Transformation results were averaged and graphed as a percentage of the highest performing reaction, T/A ligation using the Blunt/TA Ligase Master Mix.

ELECTROLIGASE™ is a trademark of New England Biolabs.

THIS WEEK

EDITORIALS

ECONOMICS Controversial study of genetic diversity and development sparks row **p.144**

WORLD VIEW Oxford mega-panel must heed lessons of other expert reviews **p.145**

ECHO CHAMBER Costa Rica bats communicate with sonar sounds **p.147**

Life stresses

It is time for sociologists and biologists to bury the hatchet and cooperate to study the effects of environmental stress on how people behave.

In the modern world, 'stress' is too often used as a catch-all word, a vague concept that bundles together a hectic pace of life and the increasing pressures that come with it. To scientists, the fuzzy notion of stress can symbolize the even fuzzier notion of the impact of an environment on an individual — one side of the classic nature–nurture debate.

Francis Galton, cousin of Charles Darwin and promulgator of his evolutionary theory, is credited with first defining the terms in this enduring conflict. "Nature is all that a man brings himself into the world; nurture is every influence which affects him after his birth," he wrote in his 1874 *English Men of Science: Their Nature and Nurture*.

That debate got fiery a century later, when the twentieth-century revolution in biology — which enabled biologists to study development and even behaviour at a molecular level — met the 1968-era social revolution. It pitted those who believed that we are determined only by our genes against those who believed we are determined only by our environment, and ignored those who pointed to the dearth of evidence either way.

Sociologists built a nurture fortress, from which they criticized what they saw as the scientific mainstream's biological determinism. They argued that the unproven and unlikely theory was dangerous because it could lead to social policies based on incorrect assumptions, such as criminals not being reformable and young minds not being vulnerable to stress. In energizing the debate, they did a service to both academia and society by keeping scientific arrogance and influence in check.

Scientists have long since abandoned any concept of biological determinism. It has now been proved beyond doubt that although our genes are fixed, their expression is highly dependent on what our environment throws at us. The current challenge is to work out precisely how environment affects our biological tissues and changes us.

Until now, analysis of the environment has been the domain of social scientists, who have elaborated the correlations between environment and health or behaviour — for example, that a deprived upbringing increases the risk of deviant behaviours in adulthood. Now, biologists are starting to render visible how one aspect of the environment — stress — leaves marks on the body (see page 161). It shortens telomeres and alters epigenetics and brain processing — and it is most potent when it occurs during brain development, a surprisingly long period of time stretching from the third trimester of pregnancy to the end of adolescence. These stress-induced changes increase vulnerability to all sorts of conditions, including psychiatric disorders and antisocial behaviour.

This rendering visible is starting to catch the attention of those who can influence social policy. For example, the influential American Academy of Pediatrics this year published a report on what it calls "toxic stress", which says that poor, or poorly coping, mothers should be cared for better while pregnant (A. S. Garner *et al. Pediatrics* **129**, e224; 2012). Last year a report commissioned by the British government covered similar ground — a development highlighted by Ilina Singh from King's College London (I. Singh *BioSocieties* 7, 308–321; 2012). Yet she is one of only a handful of sociologists to notice that a new area of potential collaboration is opening up without their input. Science is indicating that stress damage can occur in the womb, and it is understandable that politicians will look for guidance on what to do about it. But right from the start, that guidance must involve sociologists, who are best placed, for example, to balance the potential benefits of increased monitoring against possible infringement of basic freedoms and rights.

Many sociologists, however, are still immured in their fortress, struggling to catch up with a debate that has shifted from nature-or-nurture to nature-and-nurture, or are unable to shake off their distrust of scientists, worrying that scientists will force them to play second fiddle in their own territory: the environment. That is a shame: both academia and society still need their engagement at many levels. Now is a perfect time for a reconciliation of the two cultures.

> **"Only a handful of sociologists are noticing that a new area is opening up without their input."**

Funding agencies can play an important part in this. They should be aware of the need for an interdisciplinary approach and encourage their review panels to assess projects with an open mind. Psychologists, too, could offer crucial insight into these efforts, and contribute their experience of how to best to exploit the lessons of neuroscience.

Sociologists have been studying human environments for decades, and have tallied the social damage that stresses such as poverty or child abuse can cause. Biologists are now in a position to benefit from their insights, although they will need to learn the language of sociology. And sociologists stand to benefit from the understanding that biology will bring to their own, vindicated, empirical research. ■

Share alike

To make progress in clinical genomics, institutions must work out how to pass on data.

Many newborn babies admitted to intensive care have genetic disorders. The symptoms can be obvious — uncontrollable seizures, spontaneously peeling skin, abnormal heartbeats — but the cause often remains unknown. That uncertainty has painful ripples: physicians have little knowledge about how to guide treatment and parents are left unsure whether to have further children.

Genome sequencing can help. Using the fastest available sequencing instrument and software designed to guide clinicians through analysis, a team at the Children's Mercy Hospital in Kansas City, Missouri, reported in *Science Translational Medicine* last week that they had used sequencing in newborns to sift for rare genetic mutations that might cause disease (C. J. Saunders *et al. Sci. Transl. Med.* **4,** 154ra135; 2012). The results were impressive. For three of the four infants, probable culprits were identified.

To reach these conclusions, the team considered not just genetic regions in the babies, but also those in their relatives and in the scientific literature. However, for sequencing to reach its medical potential, researchers must be able to access even more genomes. Each person has millions of genetic variants — or sequences that differ from the human reference genome — making it hard to find those that might affect health. The key is to locate variants that recur in people with similar illnesses.

With so many sequencing projects under way, clinicians are always eager to know whether a variant has been observed in patients at other institutions. Analysis tools are available to help (see page 157). Yet there is currently no quick, reliable or convenient way to spread this information.

Data sharing through scientific publication has fuelled an impressive collection of databases that reveal frequencies of common variants. When variants or genes have been associated with disease, those results are also deposited in databases. ClinVar, a database from the US National Institutes of Health (NIH), for instance, gathers health-related genetic variations from the literature. And the NIH has set money aside to create a separate resource for clinically relevant genetic variants: essentially a curated database of variants for which some sort of clinical action is advised.

These are valuable efforts, but are inherently limited. Publication is too slow, and data collected about many variants will never be published. Researchers need to be able to query not just variants in the literature, but also those that have been found in other patients but not reported.

This January, an advisory group to the UK Department of Health said that the country should create a centralized facility to store genomic data to improve treatments and diagnoses. However, in the United States, where many sequencing projects are based, regulations about sharing patient data will make setting up a centralized repository more difficult.

One option would be to give patients their own sequenced genome data, letting them deposit it where they choose. Already, 23andMe, a consumer-genetics company in Mountain View, California, has used data and DNA supplied by its customers to discover (and, controversially, to patent) disease-associated variants.

Another option is for medical-research institutions to agree on ways to share information with each other. Rather than transferring a full medical record, for example, a researcher at one institute could learn whether a variant had been observed in other people and, if so, what diseases they had. For this to work, new technology and shared platforms would have to be developed.

> ***"There is currently no quick, reliable or convenient way to spread genomic information."***

There are other problems. Sequencing data are imperfect. High-throughput sequencing technology sometimes overlooks variants or makes other errors. Downstream issues are rife: are the benefits of sequencing worth the costs? Can the information be protected? How accurate are conclusions? Should information not related to the immediate medical question be shared with patients, even if a diagnosis is uncertain or no treatment is available?

All of these questions would be easier to answer if genomics data could provide more certainty. Yet to achieve this, researchers must look at the genomes and health information of more people. At this stage, the best path forward remains unclear. But for genomics to advance, the community must communicate. Institutions must consider not just what is best for their particular situation, but also what is best for the broader community. A good place to start is for staff at genomics centres and hospitals to meet in person, to share experiences and best practices. ■

Fighting chance

Collaboration between geneticists and economists has the potential to bear fruit.

One side is accused of supporting ethnic cleansing; the other of being intellectually naive. Does that sound like the beginning of a fruitful collaboration? Perhaps not, but read on.

As we report on page 154, an increasingly bitter spat has emerged between geneticists and economists over a paper that links a country's genetic diversity to its economic development.

At its heart, the argument boils down to cold statistics and methodological differences. A team of prominent geneticists and anthropologists at Harvard University in Cambridge, Massachusetts, says that the paper's economist authors did not properly account for historical and cultural connections between genetically similar countries, so correlations are mistaken for cause.

The work is part of an emerging trend to blend economics with genetics. Daniel Benjamin, an economist at Cornell University in Ithaca, New York, who is trying to identify the genetic basis for economically relevant traits such as risk aversion, is among those who say that the combination has yet to prove its worth. Nonetheless, he and others assert that understanding how genetics influences individual and international economies has the potential to inform policy.

For this to happen, both sides must take seriously the standards, methodology and history of the other. Geneticists have spent years grappling with the difficulties of getting useful information out of genomes. They have made mistakes, and learned from them, and it is naive for social scientists to think that they are immune from these errors, or that they can learn all they need to quickly. Benjamin says that nearly every study that links individual economic traits with specific genetic variants, for example, is riddled with false positives.

Social scientists should also remember that human geneticists bear the historical scars of eugenics, and more recent accusations of insensitivity to indigenous populations. Any whiff of biological determinism will draw a strong response.

Geneticists, for their part, should acknowledge that quantitative social scientists are experts in measuring human behaviour, both individual and collective. An entire subfield of economics, called econometrics, exists to make sense of data that are just as seemingly random as the string of As, Ts, Cs and Gs that comprises a genome. Moreover, many of the statistical methods that economists now use have their roots in the work of early-twentieth-century geneticists. Closer collaboration between the two fields could unlock the knowledge and expertise of social scientists, enabling them to draw conclusions that geneticists would never have conceived.

One hopeful model is the Social Science Genetics Association Consortium, a collaboration between social scientists, geneticists and epidemiologists that aims to bring more rigour to the search for the genetic basis of economic and other behavioural traits. In addition to combining the expertise of scientists in disparate fields, the consortium also has access to dozens of cohorts, encompassing more than 100,000 people.

One of its first positive results — a genetic variant linked to educational attainment in some 2,000 Icelanders — could not be replicated in other populations, raising questions over whether it is real. But the group's expertise and infrastructure gives it a chance of finding genuine links that will hopefully see geneticists working on follow-up studies, rather than writing angry letters. ■

➲ NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv

MERCK MILLIPORE

Versatility in a complete package.

MILLIPLEX®MAP magnetic bead-based immunoassay kits & MAGPIX® instruments.

Research can change direction in a minute—to keep the pace, you need a versatile assay platform.

With our analytically validated MILLIPLEX® MAP kits, using Luminex® xMAP® technology, measure one analyte per sample or fifty, in research areas including neuroscience, toxicity, oncology—using preconfigured or customized panels.

And, the versatility doesn't stop there—we're the original multiplex partner offering the full range of multiplex instruments, including MAGPIX®, the newest, most flexible, small-footprint instrument for the user-friendly xMAP® platform.

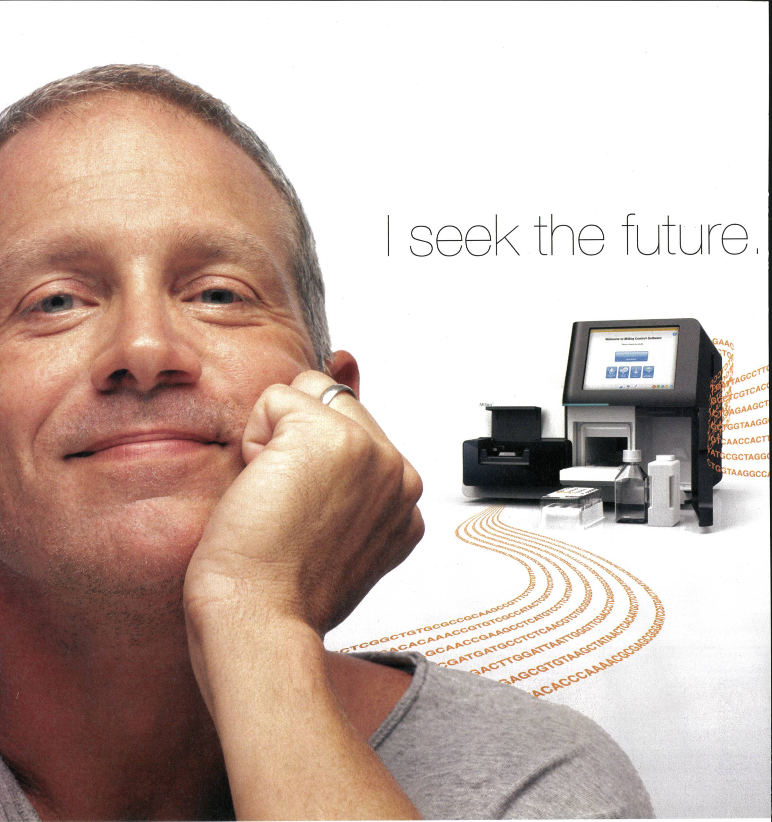
Get the complete package at
www.millipore.com/magbeads

MILLIPLEX® MAP
with Luminex® xMAP® Technology

Merck Millipore is a division of MERCK

Merck Millipore, the M logo and MILLIPLEX are trademarks of Merck KGaA, Darmstadt, Germany. Luminex, xMAP and MAGPIX are registered trademarks of Luminex Corporation. © 2012 EMD Millipore Corporation, Billerica, MA USA. All rights reserved.

THE POWER INSIDE Luminex
xMAP TECHNOLOGY Because you want to know more


I seek the future.
Welcome to MiSeq Control Software
MiSeq

MiSeq. Next-generation sequencing for all you seek.

You want amazing accuracy and performance on your benchtop. Illumina delivers—yet again. The MiSeq Personal Sequencer is the only fully integrated, truly end-to-end benchtop solution around. It's just one more example of why Illumina solutions generate a remarkable 90% of all the world's sequencing data. Discover what's possible.
www.illumina.com/iseek

illumina®

WORLD VIEW

A personal take on events

J. SUTTON-HIBBERT

Expert assessments can work if lessons are learned

For the Oxford Martin Commission on Future Generations to be heard, it should learn from previous mistakes, argues **Ehsan Masood**.

Perhaps it's just me, but I find something satisfying in watching a group of experts pore over a problem and then try to come to a collective conclusion. I started my science-journalism career covering the meetings and reports of the Intergovernmental Panel on Climate Change (IPCC) in the mid-1990s. I've tried it from the other side too — helping to organize assessments of biotechnology policy in developing countries and of the state of innovation in the Islamic world.

But expert assessments are dogged by problems, and I had begun to think that there was a growing reluctance to attempt them. That was until the University of Oxford, UK, announced in early September that they are taking another throw of the dice.

Ian Goldin, an expert in globalization and director of the university's Oxford Martin School, and Pascal Lamy, director-general of the World Trade Organization (WTO) in Geneva, Switzerland, have put together a panel of 20 experts and handed them what may be the mother of all remits.

They have asked them to look into why issues such as tackling climate change or reducing trade barriers remain unresolved, and to make their recommendations to policy-makers by the middle of next year. The grandly named Oxford Martin Commission on Future Generations promises to differ from previous expert assessments because it will also interrogate how global issues are connected to each other.

The panel includes Brazil's former foreign-affairs minister, Luiz Lampreia; the chair of South Africa's planning commission, Trevor Manuel; UK Astronomer Royal, Martin Rees; and the economist, Nicholas Stern. Seeing a list of such eminent names, it is tempting to assume that the real work will be done by a talented but overworked secretariat. But Goldin promises that his panel members are not there to decorate the cover of the final report: they are there to generate ideas. In other words, to work.

Goldin and Lamy have put a lot of thought into compiling their list. Alongside Lamy, for example, there is former World Bank president Robert Zoellick, who was also the US representative on the WTO at the same time that Lamy represented Europe as commissioner for trade. The pair, free from the burdens of representing Europe and the United States in trade talks, should be able to draw on their experiences to guide their successors.

There is also Julia Marton-Lefèvre, director-general of the International Union for Conservation of Nature in Gland, Switzerland, who will no doubt inject some environmental caution should the free-traders get too excited.

So will the Oxford Martin Commission succeed where others have failed?

To do so, it will need to learn lessons from previous assessments.

➲ NATURE.COM
Discuss this article online at:
go.nature.com/5hliq3

The first of these is about the nature of its endeavour. 'Big tent' assessments seek consensus among natural foes, but in doing so, they carry risks.

Goldin will know this well, having been at the World Bank during much of the International Assessment of Agricultural Knowledge, Science and Technology for Development, which started in 2002 and lasted for six years. An IPCC-style assessment for food with some 4,000 participants, it attempted to bridge the interests of industry with those of campaign groups, particularly over the role for genetic modification in agriculture. The biotechnology firms Monsanto and Syngenta walked out shortly before the launch of the final report.

The Oxford Martin group of 20 is a more manageable number of people, many of whom (although coming from different perspectives) are friends or know each other professionally. One might hope, then, that they will be able to resolve disagreements.

> 'BIG TENT' **ASSESSMENTS** SEEK CONSENSUS AMONG NATURAL FOES, BUT THEY CARRY **RISKS.**

The second lesson is the importance of communicating using plain language. The final report of the mighty Millennium Ecosystem Assessment in 2005, for instance, was almost impenetrable to those of us in the media, and also to policy-makers. Unusually, Goldin's expert panel features two journalists, Ariana Huffington of *The Huffington Post* and Lionel Barber of *The Financial Times*.

But I still see difficulties ahead: in particular, the tension between accessibility and providing usable knowledge. Goldin and Lamy have chosen hard problems, the solutions of which are best understood and applied by experts. Can we expect a final report rooted in interdisciplinarity, and aimed at general readers, to be taken seriously by these specialists?

Another problem is that of authority. In the heyday of global assessments, a group of experts could tell the rest of us what to think and how to behave and we would at least take note. That is an unfashionable approach today, yet nobody seems to have told the Oxford Commission.

It isn't too late. The Commission could webcast its meetings or create a rolling blog or wiki into which people can feed ideas and comment. The end result would be worth the effort, while reassuring many that this is not just an elite-level exercise.

The Oxford Commissioners must also avoid the trap of focusing exclusively on the final report. Many of the policy processes they want to fix have broken down because of a lack of trust. One thing that I've learned is that to involve your audience in the process of gathering knowledge is often as important as the eventual conclusion. ■

Ehsan Masood *is editor of* Research Fortnight *and co-author with Daniel Schaffer of* Dry: Life Without Water. *He is based in London.*
e-mail: ehm@researchresearch.com

RESEARCH HIGHLIGHTS

Selections from the scientific literature

REPRODUCTIVE BIOLOGY

Mouse eggs from stem cells

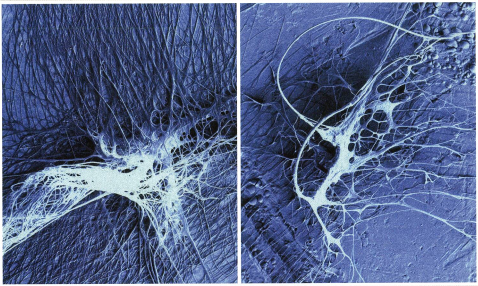
Researchers in Japan have turned mouse stem cells into egg cells that, in turn, gave rise to healthy offspring.

Katsuhiko Hayashi, Mitinori Saitou and their colleagues at Kyoto University added signalling molecules to female mouse embryonic stem cells, and to female 'reprogrammed' or induced pluripotent stem cells. This process transformed the stem cells into primordial germ cells — precursors to eggs. Culturing the germ cells with embryonic ovary tissue gave ovary-like structures, and after four weeks, the precursors developed into egg cells. The team fertilized the eggs *in vitro* and transplanted the embryos into foster mothers; the resulting offspring grew up to be fertile themselves.

The work could ultimately yield insights into potential treatments for infertility, the authors say.

Science **http://dx.doi.org/10.1126/science.1226889 (2012)**

For a longer story on this research, see http://go.nature.com/xuyljd



V. SAHNI

BIOMATERIALS

Different webs snag different prey

A common house spider uses two different types of silk structure, or attachment disc, to anchor its cobwebs — one for webs that ensnare walking insects and another for those intended to trap flying prey.

'Gumfoot' silk threads are connected to the ground and can easily release and yank walking prey into silk nets; 'scaffolding' silk is more securely attached to substrates and is used to catch flying insects. Ali Dhinojwala and his team at the University of Akron, Ohio, used electron microscopy to examine cobwebs made by the spider *Achaearanea tepidariorum,* and found that each of these webs is anchored using a different type of attachment disc. Gumfoot discs (pictured right) consist of thin arms, whereas scaffolding discs (left) have thicker fibres with many more attachment points, and stick to surfaces 10 times more strongly than gumfoot discs.

Synthetic discs made from nylon strands and tape strips displayed similar properties to their natural counterparts.

Nature Commun. **http://dx.doi.org/10.1038/ncomms2099 (2012)**

ECOLOGY

Contemplating a coral comeback

AUSCAPE/UIG/GETTY

The Great Barrier Reef (**pictured**) has lost about half its coral over the past 27 years, according to data from more than 2,000 surveys of 214 reefs.

But addressing local threats could tip the balance towards an increase in coral cover in the short term, even in the face of continuing climate change, researchers say.

Glenn De'ath of the Australian Institute of Marine Science in Townsville and his colleagues analysed the survey data and found that the mean percentage of coral cover declined from 28% to 13.8% between 1985 and 2012. Typhoons and outbreaks of the native, coral-eating crown-of-thorns starfish (*Acanthaster planci*) accounted for about 48% and 42% of the decline respectively. Bleaching related to climate change was responsible for 10%.

If severe starfish outbreaks could be prevented, perhaps by reducing nutrient run-off to the area, coral cover could recover at a rate of up to 1% per year, the researchers say. However, stabilizing global temperatures will be key to the long-term success of this strategy.

Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1208909109 (2012)**

CANCER BIOLOGY

A genome-wide 'on' switch

A common cancer gene works not by implementing unique gene-expression patterns as previously thought, but rather by boosting the expression of almost all active genes, according to work from two independent groups.

Increased expression of the gene *MYC* is often seen in human cancer, and generally bodes ill for prognosis. Richard Young at the Massachusetts Institute of Technology in Cambridge and his colleagues studied the encoded protein, c-Myc, in human tumour cells. The researchers found that c-Myc accumulates at the regulatory sequences of active genes and boosts their transcription into RNA, amplifying the gene-expression program already in place.

Meanwhile, Keji Zhao and David Levens at the National Institutes of Health in Bethesda, Maryland, and their colleagues analysed normal mouse cells. They showed that c-Myc did not induce gene expression itself, but rather amplified the expression of nearly all active genes.

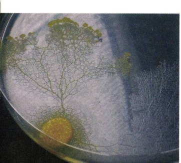
This mechanism could help to explain the diverse effects of c-Myc in different cancers.
Cell **151, 56–67; 68–79 (2012)**

ORGANISMAL BIOLOGY

How slime moulds keep track

Sticky trails of sugar and protein secreted by slime moulds help the single-celled organisms to find their way, by marking where they have previously travelled.

Christopher Reid at the University of Sydney in Australia and his team found that the slime mould *Physarum polycephalum* (**pictured**) — which senses and migrates towards chemical cues such as glucose — avoids its own secretions if possible, preferring to seek food in unexplored areas. When researchers used a U-shaped barrier to trap the creature en route to a glucose snack, 96% of organisms found their way out within 120 hours, mostly by trying untested territory. However, when the entire environment was coated in slime, only 33% of slime moulds reached the goal in that time, and those that did travelled longer distances than organisms in a slime-free setting.



A. DUSSUTOUR

Slime moulds use their trails to navigate complex environments more efficiently, the team concludes.
Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1215037109 (2012)**

EARTH SCIENCE

Planetary wanderlust

Earth's continents seem to be drifting westward by around 0.2 degrees every million years. This is a sign that the planet is experiencing true polar wander — the rotation of Earth's solid outer layer relative to its spin axis.

To distinguish true polar wander from the overlapping motion of individual tectonic plates, Pavel Doubrovine and his colleagues at the University of Oslo defined a new global reference frame, which they consider to be accurate for the past 120 million years. To do so, the researchers used the computed tracks and present positions of volcanic hot spots in the Pacific Ocean and the Indo-Atlantic hemisphere. The rate of true polar wander has increased systematically over the past 40 million years, the team found.
J. Geophys. Res. **http://dx.doi.org/10.1029/2011JB009072 (2012)**

FISHERIES

Questioning tuna marine reserves

Closing off areas of the Pacific Ocean to bigeye-tuna fishing might not be the most effective conservation measure, a modelling study suggests.

Pacific populations of bigeye tuna (*Thunnus obesus*) are being threatened by both 'purse seine' fishing, which involves catching large schools of fish with nets, and longline fishing. This led authorities to close two areas to purse seine fishing in 2009, but the effects of these closures have never been evaluated. John Sibert at the University of Hawaii in Honolulu and his colleagues used a model of tuna population dynamics to simulate the effects of different fishery-management practices on tuna biomass.

The researchers found that closing high-seas enclaves in the western central Pacific Ocean to purse seine fishing had little effect on the fishes' biomass. Reducing longline fishing in spawning areas, and prohibiting the use of fish-aggregating devices that increase the incidental catch of bigeye tuna, seem to be more efficient ways of maintaining tuna populations.
Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1209468109 (2012)**

COMMUNITY CHOICE

The most viewed papers in science

STATISTICS

In sport, winners do take all

HIGHLY READ
on www.njp.org
in September

The distribution of scores and prize money in sport is governed by the same power laws that determine the distribution of city size or wealth: a few top-ranked players and teams accumulate the vast majority of the prizes.

Wei Li of Central China Normal University in Wuhan and his colleagues came to this conclusion after studying ranking data from 12 sports, including golf, tennis and snooker. Detailed analysis of the tennis data revealed that when any two players compete, the probability that the higher-ranked player will win is related to the difference in the individuals' rankings. Moreover, the data from all 12 sports seem to follow the Pareto principle, or 80–20 rule: 20% of the players accumulate 80% of the points and money.
N. J. Phys. 14, **093038 (2012)**

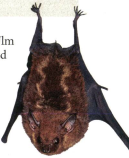
ANIMAL BEHAVIOUR

Echolocation for communication

Animals such as bats use echolocation as a form of sonar to find food at night, but they might also use it to communicate.

Mirjam Knörnschild at the University of Ulm in Germany and her colleagues recorded and analysed the echolocation calls of a social species of bat, *Saccopteryx bilineata* (**pictured**), at three sites in Costa Rica. The researchers found that roosting males seem to detect the echolocation calls of an approaching bat from at least five metres away. In response to an incoming male, the bats emitted aggressive vocalizations suggestive of territorial defence. If the approaching bat was female, however, the males responded with courtship songs.

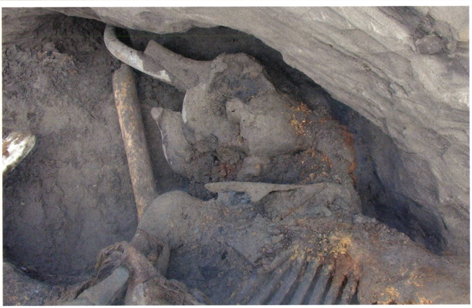
The males must be using echolocation, the authors conclude, because in low-light conditions at a distance of at least five metres, neither visual nor odour cues could provide the roosting bats with information about the sex of their visitor.
Proc. R. Soc. B **http://dx.doi.org/10.1098/rspb.2012.1995 (2012)**



M. HARVEY/GETTY

➲ NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*



PHOTAS/TASS/PA

Mammoth unearthed from Siberian mud

A remarkably well-preserved 30,000-year-old mammoth was revealed on 4 October, after an 11-year-old-boy had spotted its limbs jutting out from the frozen mud of the Taymyr peninsula in northern Russia earlier this year. Researchers excavated the 500-kilogram carcass from the Siberian tundra in September, and it was formally identified earlier this month. Researchers at the Zoology Institute in St Petersburg said that its DNA was badly damaged, probably making it useless for cloning (see *Nature* **456**, 310–314; 2008).

POLICY

GM study slammed

A study claiming that rats fed Monsanto's genetically modified NK603 maize (corn) or its companion glyphosate-based herbicide, Roundup, showed increased incidences of cancer (G.-E. Séralini *et al. Food Chem. Toxicol.* http://doi.org/jgq; 2012) has been roundly criticized by the European Food Safety Authority. On 4 October, the agency called on the study's authors to share more of their data and said that it was "presently unable to regard the authors' conclusions as scientifically sound". See page 158 for more.

Nuclear safety

Hundreds of safety upgrades are needed at European nuclear reactors, according to an analysis of the continent's power plants. Following the March 2011 Fukushima nuclear crisis in Japan, the European Commission ran safety tests that included all 145 reactors at nuclear power plants in the European Union. A report on these 'stress tests', published on 4 October, found that "practically all" nuclear plants need safety improvements. Problems ranged from the lack of back-up control rooms to substandard risk assessments. See go.nature.com/cbbiah for more.

Fracking furore

The US Environmental Protection Agency (EPA) said last week that data provided by the US Geological Survey were consistent with its December 2011 finding that the use of hydraulic fracturing — or 'fracking' — to extract natural gas had contaminated groundwater near Pavillion, Wyoming. An independent analysis commissioned by environmental groups and published on 3 October supported the EPA's assessment, but industry officials continue to question the source of the contamination. See go.nature.com/xzorhp for more.

UK funding boost

The UK government will add £200 million (US$321 million) to a fund to promote research partnerships between universities and industry, which the government says has been oversubscribed since it launched with £100 million in June. The announcement on 8 October was welcomed by scientists, although the Campaign for Science and Engineering, a London-based lobby group, noted that science investment is yet to recover from budgets slashed in 2010.

BUSINESS

Chemistry lawsuit

The American Chemical Society, the world's largest scientific society, announced on 5 October that it would pay US$22.6 million to Leadscope, a chemical-information company in Columbus, Ohio, to settle a ten-year lawsuit. The Ohio Supreme Court had found on 18 September that the society's 2002 lawsuit against Leadscope was "objectively baseless" (see *Nature* **489,** 482–483; 2012) and had been filed unfairly to crush a competitor to its Chemical Abstracts Service, which brings in some $300 million a year to the non-profit society.

Dystrophy drug hope

The experimental drug eteplirsen may help patients with Duchenne muscular dystrophy (DMD), a clinical trial of 12 boys with the condition reported on 3 October. Scientists at biotech firm Sarepta Therapeutics in Cambridge, Massachusetts, which makes the drug, revealed that four boys who had taken a high dose of eteplirsen for nearly a year were able to walk an average of 21 metres farther in six minutes than at the start of the trial. (Those on a placebo showed a decline of 68 metres.) The company plans to file for regulatory approval with the US Food and Drug Administration. About 1 in 3,600 boys develops DMD, which is caused by mutations in a gene on the X chromosome and eventually leads to paralysis and death.

Private spaceflight

California firm SpaceX launched its first mission to resupply the International Space Station on 7 October, a milestone in commercial spaceflight. As *Nature* went to press, the Dragon craft was due to dock with the space station on 10 October. The launch saw one engine fail, but the craft reached orbit. SpaceX, based in Hawthorne, said that, with nine engines, its Falcon 9 rocket was designed to handle such a problem. See go.nature.com/rvdn4f for more.

PEOPLE

Fraud funding probe

HOLLANDSE HOOGTE/EYEVINE

Diederik Stapel, a social psychologist formerly based at Tilburg University in the Netherlands, is under investigation by Dutch authorities, according to newspaper reports. Stapel resigned from his post last year after he was found to have fabricated data in some 30 published papers (see *Nature* **479,** 15; 2011). The new investigation will seek to determine whether Stapel defrauded the government by misappropriating grant money.

Nobel prizes

This year's Nobel Prize in Physiology or Medicine went to stem-cell experts John Gurdon and Shinya Yamanaka, for their work on reprogramming mature cells into their embryonic state. The physics prize was won by Serge Haroche and David Wineland, for their experiments in quantum optics. See pages 151 and 152 for more. *Nature* went to press before the chemistry prize was awarded, but full details will be available at go.nature.com/5yjkul.

Genius grants

The MacArthur Foundation, headquartered in Chicago, Illinois, selected its fellows for 2012 on 1 October. Fellowships are worth US$500,000 over 5 years, and this year, 10 out of 23 recipients are scientists, including microbiologist Sarkis Mazmanian, astronomer Olivier Guyon and marine ecologist Nancy Rabalais. The awards, popularly known as genius grants, come with no strings attached as to how the money is spent. See go.nature.com/ru2vgy for more.

EVENTS

Chemical spill

South Korea's government has designated the area around a chemical spill in the southeastern city of Gumi a special disaster zone. The 8 October announcement came nearly 2 weeks after an explosion at the Hube Globe chemical plant released some 8 tonnes of hydrofluoric acid. Five workers were killed, according to the Yonhap news agency. More than 3,000 people have since been treated after inhaling acid fumes, and the leak has damaged crops and livestock. See go.nature.com/yqjpnf for more.

RESEARCH

Telescope array

One of the world's most powerful radio-telescope arrays, the Australian Square Kilometre Array Pathfinder, was officially opened on 5 October at the Murchison Radio-astronomy Observatory in Western Australia. Composed of 36 antennas, each 12 metres in diameter, the telescope will map black holes and take a census of local galaxies, as well as testing out technology for a larger project in which it is due to be involved: the Square Kilometre Array, split between Australia and South Africa.

A year in space

Two astronauts — one American and one Russian — will stay on the International Space Station for an entire year in a mission beginning in spring 2015, NASA said on 5 October. Space-station missions are usually restricted to six months. The mission will collect more data about how humans react to long stays in space. But one year is not a record: Russian cosmonaut Valery Polyakov spent 437 days in space on the Mir space station in 1994–95.

COMING UP

13–17 OCTOBER
The Society for Neuroscience meets in New Orleans, Louisiana. Featured topics include the changing ecosystem of global neuroscience, with collaborative efforts and 'big data' coming to the fore.
www.sfn.org/am2012

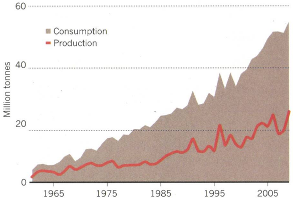
14–19 OCTOBER
New results from the Curiosity rover on Mars, and from the Kepler mission searching for extrasolar planets, are announced at the meeting of the American Astronomical Society's Division for Planetary Sciences in Reno, Nevada.
www.psi.edu/dps12

TREND WATCH

An analysis of agricultural potential in 12 African countries, released on 9 October, suggests that farmers are making use of just 10–25% of the land where wheat can be grown profitably without irrigation. Africa imports wheat because production lags behind consumption (see chart), but on the basis of modelling by the International Maize and Wheat Improvement Center in Texcoco, Mexico, the report suggests that farmers could boost domestic yields even in the face of global warming.

AFRICAN WHEAT PRODUCTION LOAFING

Wheat consumption in Africa has increasingly outstripped domestic production of the grain over the past 50 years.



SOURCE: CIMMYT/FAOSTAT

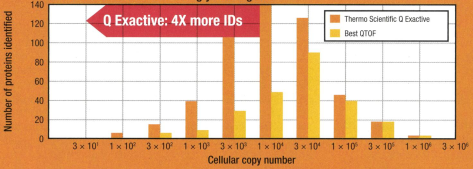
NATURE.COM
For daily news updates see:
www.nature.com/news

When it comes to scan rates, the faster the better.

Think again!

Thermo Scientific Q Exactive identified more low abundance proteins than QTOF.

Proteins identified from 10 ng yeast digest*



Exactive Plus LC-MS

Q Exactive LC-MS/MS

LTQ Orbitrap XL Hybrid FTMS

Orbitrap Velos Pro Hybrid MS

Orbitrap Elite Hybrid MS

Some manufacturers will try to convince you that you need to scan as fast as possible to obtain the most data from complex mixtures. The reality? Very fast scanning often leads to poor data quality. **Thermo Scientific Orbitrap mass spectrometers** deliver the highest effective scan rate for consistent, high quality spectra to access low-abundance proteins, making them the most desirable choice for complex samples. Learn about Orbitrap productivity and see how much more you can discover.

• visit haste makes waste at thermoscientific.com/scanspeed

* Hao Z, Zhang Y, Eliuk S, Blethrow J, Horn D, Zabrouskov V, Kellmann M, and Huhmer A. A Quadrupole-Orbitrap Hybrid Mass Spectrometer Offers Highest Benchtop Performance for In-Depth Analysis of Complex Proteomes; Thermo Scientific Application Note 552; April 2012.

Thermo
SCIENTIFIC

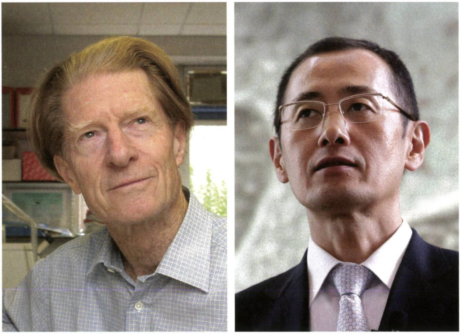
NEWS IN FOCUS

FORENSIC SCIENCE Boston crime-lab scandal exposes system overload **p.153**

SOCIAL SCIENCE Firestorm over claim that genetic diversity predicts national wealth **p.154**

BIOTECHNOLOGY Regulators assail study linking GM maize and cancer **p.158**

NEUROSCIENCE How stress scars the brain — and what can protect it **p.161**



J. PLAYER/REX FEATURES; AFLO/REX FEATURES

John Gurdon (left) and Shinya Yamanaka showed how to reprogram cells into their embryonic states.

NOBEL PRIZE

Cell rewind wins medicine Nobel

Researchers awarded prestigious prize for their work on reprogramming mature cells to a pluripotent state.

BY ALISON ABBOTT

The discovery that cells can be reprogrammed to an embryonic-like state has won this year's Nobel Prize in Physiology or Medicine for two leading lights of stem-cell research: John Gurdon and Shinya Yamanaka.

Reprogrammed cells regain pluripotency, the potential to differentiate into many mature cell types. Many researchers hope that cells created in this way will eventually be used in regenerative medicine, providing replacement tissue for damaged or diseased organs. The field has become one of the hottest in biology, but the prizewinners' discoveries were not without controversy when they were made.

Gurdon, who is based at the Gurdon Institute in Cambridge, UK, was the first person to demonstrate that cells could be reprogrammed, in work published 50 years ago[1]. At the time, scientists believed that cellular specialization was a one-way process that could not be reversed. Gurdon overturned that dogma by removing the nucleus from a frog egg cell and replacing it with the nucleus from a tadpole's intestinal cell. Remarkably, the process was able to turn back the cellular clock of the substitute nucleus. Although it had already committed to specialization, inside the egg cell it acted like an egg's nucleus and directed the development of a normal tadpole.

Gurdon was a graduate student at the University of Oxford, UK, when he did the work. He received his doctorate in 1960 and went on to do a postdoc at the California Institute of Technology in Pasadena, leaving his frogs in Europe. He did not publish the research until two years after he got his PhD, once he was sure that the animals had matured healthily. "I was a graduate student flying in the face of [established] knowledge," he says. "There was a lot of scepticism."

Mammalian cells did not prove as amenable to this process, known as cloning by nuclear transfer, as frog cells. It was nearly 35 years before the first cloned mammal — Dolly the sheep — was born, in 1996. Dolly was the only live birth from 277 attempts, and mammalian cloning remained a hit-and-miss affair.

Scientists were desperate to improve the efficiency of the system and to understand the exact molecular process involved. That is where Shinya Yamanaka of Kyoto University, Japan, made his mark. Yamanaka — who was born the year that Gurdon published his formative paper — used cultured mouse cells to identify the genes that kept embryonic cells immature, and then tested whether any of these genes could reprogram mature cells to make them pluripotent.

In the mid-2000s, the stem-cell community knew that Yamanaka was close. "I remember when he presented the data at a 2006 Keystone symposium," says Cédric Blanpain, a stem-cell biologist at the Free University of Brussels. "At that time he didn't name them and everyone was betting what these magic factors could be."

A few months later, attendees at the 2006 meeting of the International Society for Stem Cell Research in Toronto, Canada, packed out Yamanaka's lecture. The audience waited in silence before he announced his surprisingly simple recipe: activating just four genes was enough to turn adult cells called fibroblasts back into pluripotent stem cells[2]. Such induced pluripotent stem (iPS) cells could then be coaxed into different types of mature cell types, including nerve and heart cells.

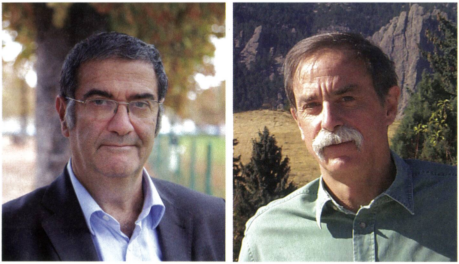
Gurdon will be 80 next year, but he continues his laboratory work on the molecular basis of reprogramming in frogs. With his mop ▶

of floppy hair and wry sense of humour, he is regarded by colleagues as a typical English gentleman, running his research institute with friendly collegiality. Gurdon sometimes notes that the honour of having an institute named after him — it was previously known as the Wellcome Trust/Cancer Research UK Institute — is usually accorded only to the dead, something colleagues can only smile at. "John is very much the active scientist," says Azim Surani, a principal investigator at the institute.

Yamanaka, who has just turned 50, is equally well-regarded by colleagues, who describe him as smartly dressed, polite and meticulous. In an interview with the Nobel Foundation, based in Stockholm, he said that the phone call informing him that he had won the award had interrupted him as he was cleaning the house. Yamanaka's research has won huge backing from the Japanese government, which now funds a large research centre for him at his university[3] and has agreed to support a stem-cell bank for clinical use[4]. Yamanaka began his career as a surgeon, but, he said, "I had no talent for it, so I decided to change my career from clinics to laboratories". "But I still feel that I am a physician — my goal, all my life, has been to bring stem-cell technologies to clinics."

Both scientists are aware that translating their discoveries into regenerative therapies will take its own time. "That's why it is so important to support basic science — it often happens that therapeutic benefit comes quite a long time after the initial discovery," Gurdon told the Nobel Foundation. ■

1. Gurdon, J. B. *J. Embryol. Exp. Morph.* **10,** 622–640 (1962).
2. Takahashi, K. & Yamanaka, S. *Cell* **126,** 663–676 (2006).
3. Cyranoski, D. *Nature* **451,** 229 (2009).
4. Cyranoski, D. *Nature* **488,** 139 (2012).



Serge Haroche (left) and David Wineland discovered ways to probe the quantum states of particles.

C. LEBEDINSKY/CNRS PHOTOTHÈQUE; NIST

NOBEL PRIZE

Physics Nobel for quantum optics

Award for methods that 'revolutionized' atomic physics.

BY GEOFF BRUMFIEL

As delicate as gossamer, the quantum properties of particles are apt to disappear as soon as physicists try to measure them. But it is possible to build a window on the quantum world to reveal these properties: and for that, Serge Haroche of the College of France, Paris, and David Wineland of the National Institute of Standards and Technology in Boulder, Colorado, have been awarded an equal share of this year's Nobel Prize in Physics.

Haroche uses atoms as a sensitive probe of light particles trapped in a cavity, whereas Wineland takes the opposite approach, using light to measure the quantum states of atoms. Both techniques have helped to investigate the fundamentals of quantum mechanics, and they are helping to develop new technologies such as quantum computers or atomic clocks of dizzying precision. News of the award came as a shock to Haroche: "I recognized the Swedish phone code. I had to sit down," he said at a press conference shortly after the announcement.

In the quantum world, particles of light and matter obey strange rules. One particle can occupy several mutually exclusive states simultaneously, for example, and groups of particles can be mysteriously connected through a process known as entanglement. But these quantum properties are hard to see: particles will show their quantum nature only in isolation, and even the slightest bump from the outside world will destroy their quantum states. That makes experiments extremely tricky, because the act of measuring itself is enough to upset the system. The techniques developed by Wineland and Haroche gave physicists a way to probe these states without destroying them.

Haroche's experiments bounce microwave photons between a pair of superconducting mirrors, and send a stream of rubidium atoms through the fog of photons. By measuring the spins of the atoms as they enter and exit the mirrored cavity, he is able to indirectly probe the quantum properties of the microwave photons inside. Progressive measurements have, for example, allowed his team to observe a photon's quantum wavefunction — which simultaneously describes all of its possible quantum states — and then monitor its collapse to a single, well-defined state[1].

Wineland's group traps beryllium ions in electric fields, and cools them with a laser that excites the ion's electrons. This sucks vibrational energy from the system, lowering the temperature[2]. Researchers can then use lasers to alter vibrations between the ions, allowing them to control the quantum interactions in the system[3]. The work is already being used to build atomic clocks with unprecedented accuracy, says Immanuel Bloch, a physicist at the Max Planck Institute for Quantum Optics in Garching, Germany. Further down the line, these techniques could be used in a quantum computer — a device that can perform calculations using the probabilistic rules of quantum mechanics.

The award is "a great choice of two people who have really contributed to the foundations of quantum physics", Bloch says. He notes that this is just the latest in a run of Nobel prizes for quantum optics. Bloch thinks that this is down to the myriad techniques, such as those of Wineland and Haroche, that are allowing researchers to isolate, study and manipulate increasingly complex quantum systems. "I think we've really seen atomic physics revolutionized," he says. ■

1. Guerlin, C. *et al. Nature* **448,** 889–893 (2007).
2. Monroe, C. *et al. Phys. Rev. Lett.* **75,** 4011–4014 (1995).
3. Monroe, C. *et al. Phys. Rev. Lett.* **75,** 4714–4717 (1995).

B. TESFAYE/AP PHOTO

Annie Dookhan, a former forensic chemist in Massachusetts, has been charged with obstructing justice.

FORENSICS

Boston scandal exposes backlog

Chemist charged with fraud shows system under pressure.

BY EUGENIE SAMUEL REICH

The numbers are shocking: 1,100 people are in jail on the basis of evidence that may have been tampered with, and 34,000 criminal cases could be affected. The fallout stems from the work of just one forensic chemist, who is alleged to have faked test results on suspected drug samples.

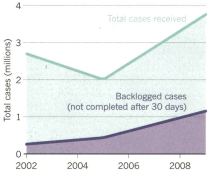
The revelations and the subsequent arrest on 28 September of Annie Dookhan, who worked at a state-run laboratory in Boston, Massachusetts, have rattled the US forensics community and raised questions about oversight across the profession. *Nature* has learned that the facility had applied for federal funding to help to clear a backlog of some 8,000 cases — a full year's work. The situation parallels that of overburdened forensic labs across the United States.

"If you think this couldn't happen in your lab, you're naive," says Robin Cotton, director of the forensic-sciences programme at Boston University and a former forensic-lab director. "Lab directors all over the country are paying attention." Cotton and others say that the affair highlights a resourcing problem that is putting pressure on US forensic-lab workers and their supervisors, who need to bring cases to court fast, but without compromising the integrity of test results.

Dookhan has pleaded not guilty to charges of obstructing justice by falsifying data in two criminal cases, and of lying under oath about her qualifications. However, a wide-ranging confession by Dookhan documented in a police report published by *The Boston Globe* implies that she falsified results on numerous occasions over the past 2–3 years. In most cases, she had guessed the nature of samples she had been asked to test; but "a few times" she had recorded the results of negative drug tests as being positive. Because Dookhan wasn't able to tell police which samples she had faked, and has said that she sometimes contaminated samples after the fact so that they would conform to her guesses, the possibility of wrongful conviction now swirls around the tens of thousands of cases on which she worked. Those who are currently serving jail terms are set to have their cases reviewed in court hearings that will begin on 15 October.

SYSTEM OVERLOAD

Demands on US forensic labs have expanded — and so has the proportion of backlogged cases.



Dookhan's role at the lab was to identify compounds found by police that were suspected to be drugs. Police interviews with her former colleagues reveal that she tested more samples in less time than any other chemist in the lab — a record of productivity that raised eyebrows and had generated complaints from co-workers. In one instance, a top official reacted to concerns by giving Dookhan a special project to try to "slow her down".

Dookhan resigned in March while facing termination proceedings. The director of analytical chemistry has been fired, the lab bureau chief has resigned and Dookhan's direct supervisor faces disciplinary proceedings. Massachusetts public-health commissioner John Auerbach has also resigned, taking responsibility for the loss of confidence in the state's forensic work as a result of the scandal.

RESOURCE PROBLEM

Forensic labs in the United States are under a great deal of pressure to keep up with increasing case loads. In Massachusetts, for example, *Nature* has found that the National Institute of Justice (NIJ) in Washington DC allocated the state police at least US$1.2 million from 2009 and 2011 to try to clear backlogs, in part by hiring an extra drug chemist and by buying new equipment. A spokeswoman for the NIJ has confirmed that some of the funding was routed to the lab in which Dookhan worked. The grants came with a requirement to provide oversight and report on any allegations or suspicions of wrongdoing that might affect the integrity of forensic science, but the NIJ says that nothing had been reported by state police by the end of 2011.

Censuses of US forensic labs show that as many as one in four requests by law enforcement for forensic work are not acted on within 30 days. That has led to an ever-increasing backlog that accounts for a growing proportion of the workload (see 'System overload'). The backlog is worst for DNA testing, which has become a major focus for groups interested in wrongful convictions. However, with drug law enforcement a priority in the US justice system, labs that test drugs are also being

▶

swamped with cases.

Interviews documented in the police report suggest that a key cause of the backlogging at the lab is a 2009 US Supreme Court decision known as *Melendez-Diaz* v. *Massachusetts*, which overturned a drug conviction because the defence had been refused permission to cross-examine the forensic scientists who tested samples seized from a suspect. The result means that defence attorneys are now more likely to call forensic scientists into court to testify. "They're spending all their time in the courtroom and not the laboratory," says Ralph Keaton, executive director of the American Association of Crime Laboratory Directors Laboratory Accreditation Board in Garner, North Carolina. "Then the backlog grows."

A bill now before the US Congress aims to improve the quality of forensic science by providing funding for research and for development of national standards. But the bill does not include funding to clear backlogs. Stephen Saloom, policy director at the Innocence Project in New York, which seeks to uncover wrongful convictions, says that even if the bill passes it won't provide an immediate remedy for cases of deliberate evidence tampering.

Dookhan's attorney, Nicolas Gordon, says that his client is not speaking to the media. Dookhan has been released on bail until a court hearing on 3 December. Gordon acknowledges allegations against his client but won't comment on their veracity. "It's a fluid situation that could change over the next few months," he notes. ■

INTERDISCIPLINARY RESEARCH

Economics and genetics meet in uneasy union

Use of population-genetic data to predict economic success sparks war of words.

BY EWEN CALLAWAY

"The invalid assumption that correlation implies cause is probably among the two or three most serious and common errors of human reasoning." Evolutionary biologist Stephen Jay Gould was referring to purported links between genetics and an individual's intelligence when he made this familiar complaint in his 1981 book *The Mismeasure of Man*.

Fast-forward three decades, and leading geneticists and anthropologists are levelling a similar charge at economics researchers who claim that a country's genetic diversity can predict the success of its economy. To critics, the economists' paper seems to suggest that a country's poverty could be the result of its citizens' genetic make-up, and the paper is attracting charges of genetic determinism, and even racism. But the economists say that they have been misunderstood, and are merely using genetics as a proxy for other factors that can drive an economy, such as history and culture. The debate holds cautionary lessons for a nascent field that blends genetics with economics, sometimes called genoeconomics. The work could have real-world pay-offs, such as helping policy-makers to set the right level of immigration to boost the economy, says Enrico Spolaore, an economist at Tufts University near Boston, Massachusetts, who has also used global genetic-diversity data in his research.

The United States has the right amount of genetic diversity to buoy its economy, claim economists.

D. ACKER/BLOOMBERG VIA GETTY

But the economists at the forefront of this field clearly need to be prepared for harsh scrutiny of their techniques and conclusions. At the centre of the storm is a 107-page paper by Oded Galor of Brown University in Providence, Rhode Island, and Quamrul Ashraf of Williams College in Williamstown, Massachusetts[1]. It has been peer-reviewed by economists and biologists, and will soon appear in *American Economic Review*, one of the most prestigious economics journals.

The paper argues that there are strong links between estimates of genetic diversity for 145 countries and per-capita incomes, even after accounting for myriad factors such as economic-based migration. High genetic diversity in a country's population is linked with greater innovation, the paper says, because diverse populations have a greater range of cognitive abilities and styles. By contrast, low genetic diversity tends to produce societies with greater interpersonal trust, because there are fewer differences between populations. Countries with intermediate levels of diversity, such as the United States, balance these factors and have the most productive economies as a result, the economists conclude.

The manuscript had been circulating on the Internet for more than two years, garnering little attention outside economics — until last month, when *Science* published a summary of the paper in its section on new research in other journals. This sparked a sharp response from a long list of prominent scientists, including geneticist David Reich of Harvard Medical School in Boston, Massachusetts, and Harvard University palaeoanthropologist Daniel Lieberman in Cambridge.

In an open letter, the group said that it is worried about the political implications of the economists' work: "the suggestion that an ideal level of genetic variation could foster economic growth and could even be engineered has the potential to be misused with frightening consequences to justify indefensible practices such as ethnic cleansing or genocide," it said.

The critics add that the economists made blunders such as treating the genetic diversity of different countries as independent data, when they are intrinsically linked by human migration and shared history. "It's a misuse of data," says Reich, which undermines the paper's main conclusions. The populations of East Asian countries share a common genetic history, and cultural practices — but the former is not necessarily responsible for the latter. "Such haphazard methods and erroneous assumptions of statistical independence could equally find a genetic cause for the use of chopsticks," the critics wrote.

They have missed the point, responds Galor, a prominent economist whose work examines the ancient origins of contemporary economic factors. "The entire criticism is based on a gross misinterpretation of our work and, in some respects, a superficial understanding of the empirical techniques employed," he says. Galor and Ashraf told *Nature* that, far from claiming that genetic diversity directly influences economic development, they are using it as a proxy for immeasurable cultural, historical and biological factors that influence economies. "Our study is not about a nature or nurture debate," says Ashraf.

"It seems like the devil is in the interpretation more than the actual application of the statistics," says Sohini Ramachandran, a population geneticist at Brown University who provided the genetic data for the study. She adds that Galor and Ashraf used estimates of genetic diversity that she and her colleagues specifically developed to overcome many of the confounding factors caused by the overlapping genetic and cultural histories of neighbouring countries.

Galor and Ashraf are not the first economists to use genetic-diversity data. Spolaore has also found that the differences in genetic diversity between countries can predict discrepancies in their level of economic development[2]. But he is clear that this is not necessarily a causal relationship: "In my view it's not genetic diversity itself that is responsible for this correlation," he says. "A lot of this could be culture."

Some say that the field needs a dose of rigour. Many studies linking genetic variation to economic traits make basic methodological errors, says Daniel Benjamin, a behavioural economist at Cornell University in Ithaca, New York. He is part of the Social Science Genetics Association Consortium, a group that brings together social scientists, epidemiologists and geneticists to improve such studies. Problems that medical geneticists have known about for years — such as those stemming from small sample sizes — crop up all too often when economists start to work with the data, he says.

> ***"Our study is not about a nature or nurture debate."***

For instance, while searching for genetic associations with factors such as happiness and income in a study of 2,349 Icelanders, Benjamin and his colleagues found a statistically significant association between educational attainment and a variant in a gene involved in breaking down a neurotransmitter molecule[3]. But the researchers could not replicate this association in three other population samples — a test for false positives that is standard practice in medical genetics — and the team now has reservations about the association. If the field is to develop fruitfully, "I think it's essential for us to have geneticists involved", says Benjamin. "We couldn't do it without their help and insight." ■ **SEE EDITORIAL P.144**

1. Ashraf, Q. & Galor, O. *Am. Econ. Rev.* (in the press).
2. Spolaore, E. & Wacziarg, W. *Q. J. Econ.* **124,** 469–529 (2009).
3. Benjamin, D. J. *et al. Annu. Rev. Econ.* **4,** 627–662 (2012).

MORE ONLINE

TOP STORY

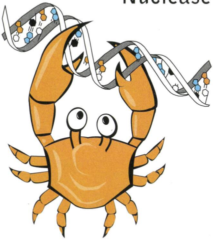
JGI/J. GRILL/GETTY

- Brain scans could help in design of reading lessons. go.nature.com/q4lbmf

MORE NEWS

- Despite conservation efforts, most whale deaths are caused by humans. go.nature.com/9sy1vz
- US states make vaccination opt-out harder. go.nature.com/rjpkv4
- Proteins remember the past to predict the future. go.nature.com/3darfm

King Crab
Duplex-Specific
Nuclease

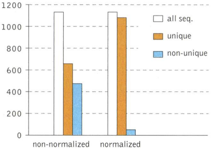


- Construction of normalized RNA-seq libraries
- Construction of normalized full-length enriched cDNA libraries
- Ribosomal cDNA depletion

Evrogen offers kits and custom services utilizing unique properties of thermostable Duplex-Specific Nuclease (DSN), purified from hepatopancreas of Red King (Kamchatka) crab.

The DSN-based normalization protocols developed by Evrogen, Illumina® and independent teams are proved to be a convenient and effective approach to remove repeated sequences from DNA mixture.

The protocols exploit nucleic acid hybridization kinetics and a strong preference of DSN for cleaving double-stranded DNA, and are now compatible with all main NGS platforms.



Transcript distribution in standard and normalized cDNA libraries from *Aplysia* neurons: white – all sequences, orange – unique sequences; blue – non-unique sequences.

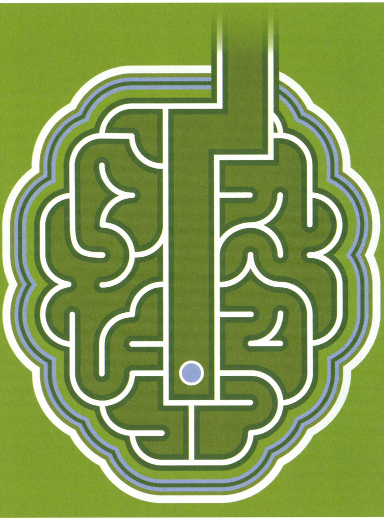
Evrogen JSC, Moscow, Russia
Tel: +7(495) 988 40 84
Fax: +7(495) 988 40 85
E-mail: evrogen@evrogen.com
Web site: www.evrogen.com

evrogen

BioResearch

Lonza

Take a Direct Route to Results



Neurobiology Research Tools for Biologically Relevant Results

Go straight to the data you need with Lonza's integrated primary cells and media, transfection technology, and qPCR arrays.

Designed specifically for neurobiology studies, our research tools will help keep you on the fast track to biologically relevant results.

Visit us to learn more about our:

- Clonetics™ Primary Cells and Media, tailored for neurobiology research
- Nucleofector™ Transfection Technology for primary neurons
- StellARray™ Disease- and Pathway-specific qPCR Arrays

The direct path to results starts with Lonza.

Lonza Research Solutions
© 2011 Lonza Walkersville, Inc. www.lonza.com/research

www.lonza.com/neurobiology

BUSINESS

Genome interpreter vies for place in clinical market

Launch of system that keeps data local aims to address privacy fears.

BY MONYA BAKER

As DNA sequencing gets faster and cheaper, clinicians are clamouring to use it. A test for malfunctioning genes might show how to treat a tumour or help to diagnose the underlying causes of a disease. But sequencing data are too complex for most clinicians to analyse, and medical institutions are wary of transferring patient data to specialists elsewhere for analysis.

A genome-interpretation company is now offering its solution: a 1-metre-tall, 275-kilogram black box that carries enough storage and processing power to analyse one genome every day, picking out mutations with potential links to disease — in theory, fast enough to inform treatment. But for some, the most important feature of the US$125,000 unit is that it is a self-contained object. In an era of cloud computing and global networks, a machine that keeps its information stubbornly local has growing appeal. "There is a tremendous worry about privacy with sharing patient data," says Martin Tolar, chief executive of Knome, the company in Cambridge, Massachusetts, that produces the device. "The institutions we approached said, 'We want to keep the system within our four walls.'"

Unveiled on 27 September, the Knome system, knoSYS 100, belongs to an emerging class of services and tools to help clinical researchers to catch up with advances in genome sequencing. That capability "has made itself available faster than we are prepared to deal with", says Vincent Funari, director of the genomics core facility at Cedars-Sinai Medical Center in Los Angeles, California.

The core of Knome's system is not hardware, but software. The machine combs through a newly sequenced genome to find places where humans vary, and annotates them with existing knowledge. This process, known as genome interpretation, can winnow down the millions of variants found in any individual's genome to a handful that might explain a disease (see 'Super sifter'). "Our goal is to take these data and say, 'For this group of patients, these are the 5–10 variants that are most likely to be implicated,'" explains Tolar. For now, the software is not meant for clinical diagnosis or medical advice. Clients include drug companies and medical centres researching how to use sequencing for clinical decisions.

➲ NATURE.COM
Read more about whole genomes in the clinic at:
go.nature.com/yv7rls

SUPER SIFTER

Genome-interpreting tools work by discarding genetic differences that look inconsequential or lack information, and focusing on a handful likely to be linked to disease. (Schematic is simplified; tools vary.)

INPUT: Sequenced genome (~6 billion base pairs)	Variants identified
Compare with reference genome and known variation	3 million to 5 million variants
Filter for variants in genes	>20,000 variants
Compare with genetic databases and genomes of other ill or healthy people to select rare variants and those associated with disease	~1,000 variants
If no known causative variant found, run through algorithms to predict variants likely to disrupt gene function	~100 variants
Use existing knowledge to find variants most likely to cause disease	~10 variants
OUTPUT: For follow-up analysis	~10 variants

Experts warn that genome interpretation is far from mature, and that its reliability depends on the quality of the sequences it analyses. Not all types of variants can be detected, and errors occur at every step before and during interpretation: in the sequencing of fragments of DNA; in matching those millions of fragments with their equivalents in the reference genome; and in detecting differences. "The interpretation of variants is absolutely dependent on accurate variant identification," says Karl Voelkerding, medical director for genomics and bioinformatics at ARUP Laboratories, an assay facility in Salt Lake City, Utah.

Then there is the challenge of working out which detected variants are relevant to disease. The protocols are imperfect and the various annotation tools access different data in different ways and so supply a variety of answers. And all annotation tools uncover many 'variants of unknown significance', about which too little is known to assess whether or not they affect a person's health. Any variant that might be used to provide a diagnosis or guide patient care must be verified independently by separate experiments.

But just organizing information into a useful form is a big step forward, says George Church, a genomicist at Harvard Medical School in Boston, Massachusetts, and co-founder of Knome. The process, he says, "is not about perfection. It's about delivering a high-quality interpretation based on current knowledge."

Knome's device may be well placed to tap a nascent clinical market in which data is preferentially kept on site, but sequencing companies are also making inroads with software that requires genomic data to be transferred elsewhere. For example, Illumina, based in San Diego, California, offers free data storage and variant identification for clients who upload sequencing data to its cloud-computing platform, which has an open programming interface. Illumina has contracts with a suite of other companies to develop data-analytic applications in the cloud. As more centres begin using sequencing data, it is expected that policies and procedures for using the cloud will mature.

David Goldstein, a genomicist at Duke Institute for Genome Sciences and Policy in Durham, North Carolina, says that although commercial genome interpreters may be valuable in the clinic, he doubts that they will make headway in the larger genomics research centres. Commercial offerings rely on standard references and techniques that can make their approach too constraining for genomics research, he says. "We've never had a 12-month stretch where we've done something the same way."

But for clinicians interested more in working out how best to apply current genomics knowledge, prepackaged software and hardware may be just the ticket. "Last year we had zero clinical institutes" as clients, says Tolar, "and this year we have 20." ■ SEE EDITORIAL P.143

BIOTECHNOLOGY

Hyped GM maize study faces growing scrutiny

Food-safety bodies slam feeding study that claims increased cancer incidence in rats.

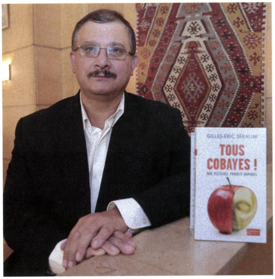
BY DECLAN BUTLER

The storm of scientific criticism over claims that a genetically modified (GM) maize causes severe disease in rats shows no signs of abating.

Gilles-Eric Séralini, a molecular biologist at the University of Caen, France, is under intense pressure to report the full data behind his team's finding that rats fed for two years with Monsanto's glyphosate-resistant NK603 maize (corn) developed many more tumours and died earlier than controls (see *Nature* **489,** 484; 2012). The study, run in collaboration with the Paris-based Committee for Research and Independent Information on Genetic Engineering (CRIIGEN), also found that rats developed tumours when their drinking water was spiked with glyphosate, the herbicide that is used with the GM maize. The findings have had a huge public impact in Europe, empowering those opposed more broadly to GM foods, and leading some politicians to call for tighter regulations or outright bans of the maize.

Last week, the European Food Safety Authority (EFSA) in Parma, Italy, and Germany's Federal Institute for Risk Assessment (BfR) in Berlin both issued initial assessments slamming the paper, bluntly asserting that its conclusions are not supported by the data presented. "The design, reporting and analysis of the study, as outlined in the paper, are inadequate," says the EFSA in a press release, adding that the paper is "of insufficient scientific quality to be considered as valid for risk assessment".

The biggest criticism from both reviews is that Séralini and his team used only ten rats of each sex in their treatment groups. That is a similar number of rats per group to that used in most previous toxicity tests of GM foods, including Missouri-based Monsanto's own tests of NK603 maize. Such regulatory tests monitor rats for 90 days, and guidelines from the Organisation for Economic Co-operation and Development (OECD) state that ten rats of each sex per group over that time span is sufficient because the rats are relatively young. But Séralini's study was over two years — almost a rat's lifespan — and for tests of this duration, the OECD recommends at least 20 rats of each sex per group for chemical-toxicity studies, and at least 50 for carcinogenicity studies.



Gilles-Eric Séralini's book describes his latest GM research.

J. DEMARTHON/AFP/GETTY

Moreover, the study used Sprague-Dawley rats, which both reviews note are prone to developing spontaneous tumours. Data provided to *Nature* by Harlan Laboratories, which supplied the rats in the study, show that only one-third of males, and less than one-half of females, live to 104 weeks. By comparison, its Han Wistar rats have greater than 70% survival at 104 weeks, and fewer tumours. OECD guidelines state that for two-year experiments, rats should have a survival rate of at least 50% at 104 weeks. If they do not, each treatment group should include even more animals — 65 or more of each sex.

"There is a high probability that the findings in relation to the tumour incidence are due to chance, given the low number of animals and the spontaneous occurrence of tumours in Sprague-Dawley rats," concludes the EFSA report. In response to the EFSA's assessment, the European Federation of Biotechnology — an umbrella body in Barcelona, Spain, that represents biotech researchers, institutes and companies across Europe — called for the study to be retracted, describing its publication as a "dangerous case of failure of the peer-review system".

Séralini argues that a battery of observations in the study reinforces his tumour-incidence and mortality claims. "Of course, this should be replicated by others, but we believe in these results," he says. He agrees that more rats would have boosted his study's statistical power, but says that he did not design the experiment to show differences in tumour incidences, because he was not expecting to find any — no previous tests on GM foods had suggested a cancer risk.

Yet Séralini has promoted the cancer results as the study's major finding, through a tightly orchestrated media offensive that began last month and included the release of a book and a film about the work. Only a select group of journalists (not including *Nature*) was given access to the embargoed paper, and each writer was required to sign a highly unusual confidentiality agreement, seen by *Nature*, which prevented them from discussing the paper with other scientists before the embargo expired.

Journalists often receive embargoed journal articles, and standard practice is to solicit independent assessments before the paper is published. The agreement for this paper, however, did not allow any disclosure and threatened a severe penalty for non-compliance: "A refund of the cost of the study of several million euros would be considered damages if the premature disclosure questioned the release of the study."

In an exceptional move, the ethics committee of the French National Centre for Scientific Research (CNRS) last week decried the public-relations offensive as inappropriate for a high-quality and objective scientific debate, and reminded researchers working on controversial topics of the need to report results responsibly to the public.

Meanwhile, Séralini says that he won't make any data available to the EFSA and the BfR until the EFSA makes public all the data underpinning its 2003 approval of NK603 maize for human consumption and animal feed. He has also criticized the EFSA, and most other detractors of his study, for alleged conflicts of interest, claiming that he is "being attacked in an extremely dishonest fashion by lobbies passing themselves off as the scientific community".

The journal that published his study, *Food and Chemical Toxicology*, said last week in a statement that it "welcomes any and all 'Letters to the Editor' that have questions and concerns about this paper". ■

GRANTS

US biologists decry funding changes

National Science Foundation policy may harm tenure seekers.

BY HELEN SHEN

“This is the farthest I think that I’ve stuck my neck out,” says Sarah Hobbie, an ecologist at the University of Minnesota in Saint Paul. Policy changes at the US National Science Foundation (NSF) are giving life scientists in some fields fewer opportunities to apply for funding, so Hobbie is now at the forefront of a campaign to persuade the NSF to rethink its approach.

In August, Hobbie garnered more than 550 signatures with an open letter to the NSF that says the changes, put in place last year, are jeopardizing the careers of young scientists. She has also surveyed the ecology community, which relies heavily on NSF funding, about what it thought was the best way forward for the agency. Of the 1,622 responses she received, only 13% favoured the NSF’s changes over an alternative devised by Hobbie and her colleagues.

At the heart of the dispute are changes to the NSF’s application process for grants in organismal and environmental biology. The two divisions account for about half of the US$700 million that the agency will spend on the life sciences in 2012. In the past, researchers had two chances to apply for funding each year. Now they have just one opportunity, in two stages. For the first stage, researchers submit a short ‘pre-proposal’ in January, which is then screened by reviewers. Researchers who clear this hurdle are invited to submit full proposals by 2 August. Grant winners are awarded funding the following year.

“With a single decision, you’ve made an abrupt change that will affect a lot of people.”

The NSF says that it expects to award roughly the same number of grants as in previous years, but the revised process will reduce the burden on grant reviewers. “The old system had reached breaking point,” says John Wingfield, who heads the biology directorate at the NSF in Arlington, Virginia. He adds that the change has “created some interest in other directorates”.

Hobbie, whose two pre-proposals were turned down this year, must now wait until January to try again for the chance to submit a full proposal in August. If approved, that project would probably not receive funding until 2014. She says that such long gaps between resubmissions have a domino effect because the delays to research and publications can be career-limiting for younger researchers striving to establish themselves and earn tenure.

“Instead of 10 or 12 opportunities to put in a proposal before tenure, it’s 5 or 6 now,” says Jim Heffernan, an ecologist in his first year of an assistant professorship at Duke University in Durham, North Carolina.

Another change, which prohibits researchers from acting as principal investigators on more than two proposals a year, is equally unpopular. “Strategically, I need to be the lead principal investigator on whatever grant I get,” says Stan Harpole, an ecologist and assistant professor at Iowa State University in Ames, who is two years from his tenure evaluation.

COMMON CONCERN

Simmering criticisms of the new policies bubbled over in August at a panel discussion during the annual meeting of the Ecological Society of America in Portland, Oregon. It was there, Hobbie says, that she realized that her colleagues shared her concerns. The open letter she sent later that month led to a private meeting with Wingfield on 10 September. A week later, the heads of 33 life-sciences departments and programme leaders from several US universities sent a letter to the NSF expressing similar worries.

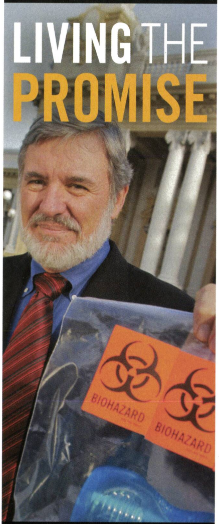
“With a single decision at the NSF, you’ve made an abrupt change that will affect a lot of people,” says Kenneth Petren, head of biological sciences at the University of Cincinnati in Ohio, who signed the second letter. “I’m most concerned younger people will decide there’s a better career path for them.”

On 3 October, Wingfield told *Nature* that he understands the concerns, but that the NSF has opted not to reverse its policy for now. Instead, he says, the agency will evaluate the changes over the next few grant cycles. Hobbie and Petren say that such monitoring should involve dialogue with affected researchers.

“Engaging with the community is the only way they’re going to assess certain types of impacts on science,” says Hobbie. “I’m quite disappointed that I haven’t heard anything in that direction.” ■

UNIVERSITY OF CALIFORNIA
UCRIVERSIDE

LIVING THE PROMISE



Real World Solutions

Legally poisoned:
UCR Prof. Carl Cranor uncovers the alarmingly high rates of toxic chemicals that we are exposed to in everyday products and pushes for better protections in public health law.

Explore more impacts >
promise.ucr.edu

Get your personal qPCR cycler

Rotor-Gene® Q

Experience high-performance qPCR at a budget-conscious price

Running
Standby
Rotor-Gene Q
QIAGEN
FAST

The Rotor-Gene Q Personal Edition 2plex and 2plex HRM® instruments.

- **Thermal uniformity for high-quality reproducible data**
- **Exceptional High-Resolution Melt technology**
- **Flexible throughput, easily adaptable to your workflow needs**
- **Robust design for minimum maintenance**
- **Free software upgrades thanks to unlimited user license**

Learn more at www.qiagen.com/rgqpersonal

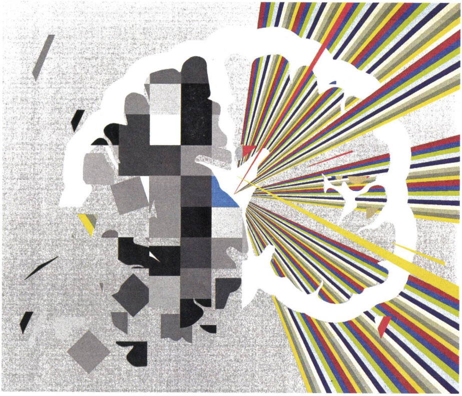
AUTORGQPersonal0712N1WW

Pricing may vary from country to country. Not available in all countries. Void where prohibited. Offer not valid for Rotor-Gene Q MDx instruments. Not valid for diagnostic applications. Offer expires December 31, 2012. Please contact your local QIAGEN sales representative or distributor for more information. Trademarks: QIAGEN®, HRM®, Rotor-Gene® (QIAGEN Group). For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual (available at www.qiagen.com or from your local distributor).

QIAGEN®

Sample & Assay Technologies

PADDY MILLS



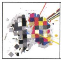
UNDER PRESSURE

A special section of *Nature* traces the emerging links between stress and mental illness.

A battlefield; an abusive parent; the ongoing struggle to make ends meet; a pile-up of unanswered e-mails — stress can take many forms. But stressful situations have something in common: they trigger reactions such as fear and a surge of hormones. These responses can be beneficial in the short term but, if severe or prolonged, can damage the mind, triggering conditions ranging from depression to post-traumatic stress disorder. That much has been known for decades; now researchers are getting to grips with how stress can alter the biology of the brain, and tip a mind into illness. Here, *Nature* takes a look at what they have learned, and at the gaps that remain.

Modern life — with more people, more pressure and more gadgets — certainly seems more stressful than life in the past, but how does it affect the mind? Some scientists are seeking answers by examining how the brains of city- and country-dwellers process stressful situations (see page 162). Others are looking for the molecular scars left by stress. Neuroscientist Eric Nestler argues that stress may influence the brain through epigenetics, mechanisms that change how genes behave (see page 171). Biologists Elizabeth Blackburn and Elissa Epel, meanwhile, suggest that stress causes chronic diseases in part by shortening telomeres, structures that cap and protect the ends of chromosomes (see page 169).

Still others hope to draw lessons from people who can bounce back from stress — even from experiences as devastating as abduction and rape. By tracing the social and biological factors that help these people to recover, the research could yield ways to make the rest of us more resilient (see page 165). ■ **SEE EDITORIAL P.143 AND CAREERS P.299**



STRESS AND RESILIENCE
The links between adversity and mental illness. **nature.com/stress**

R. MANSI/GETTY

URBAN DECAY

SCIENTISTS ARE TESTING THE IDEA THAT THE STRESS OF MODERN CITY LIFE IS A BREEDING GROUND FOR PSYCHOSIS.

BY ALISON ABBOTT

STRESS AND RESILIENCE
The links between adversity and mental illness. nature.com/stress

In 1965, health authorities in Camberwell, a bustling quarter of London's southward sprawl, began an unusual tally. They started to keep case records for every person in the area who was diagnosed with schizophrenia, depression, bipolar disorder or any other psychiatric condition. Decades later, when psychiatrists looked back across the data, they saw a surprising trend: the incidence of schizophrenia had more or less doubled, from around 11 per 100,000 inhabitants per year in 1965 to 23 per 100,000 in 1997 — a period when there was no such rise in the general population (J. Boydell *et al. Br. J. Psychiatry* **182,** 45–49; 2003). The result raised a question in many researchers' minds: could the stress of city life be increasing the risk of schizophrenia and other mental-health disorders?

The question is an urgent one. Back in 1950, less than one-third of the world's population lived in cities. Now, lured by the prospect of work and opportunity, more than half do. Mental illnesses already comprise the world's biggest disease burden after infectious diseases and, although global statistics do not yet show any major increase in incidence, the cost is rising. In Germany, the number of sick days taken for psychiatric ailments doubled between 2000 and 2010; in North America, up to 40% of disability claims for work absence are related to depression, according to some estimates. "It seems that cities may be making us sick," says Jane Boydell at the Institute of Psychiatry in London, who led the Camberwell study.

Anecdotally, the link between cities, stress and mental health makes sense. Psychiatrists know that stress can trigger mental disorders — and modern city life is widely perceived as stressful. City dwellers typically face more noise, more crime, more slums and more people jostling on the streets than do those outside urban areas. Those who have jobs complain of growing demands on them in the workplace, where they are expected to do much more in less time.

But the idea has not been widely tested. It is difficult to study whether something as complex as a 'city environment' has an impact on the brain. To complicate matters, many growing cities include immigrant populations, which already have an increased risk of psychiatric disease associated with social isolation.

Now, a few scientists are tackling the question head on, using functional brain imaging and digital monitoring to see how people living in cities and rural areas differ in the way that their brains process stressful situations. "Yes, city-stress is a big, messy concept, but I believed it should be possible to at least see if brains of city-dwellers looked somehow different," says Andreas Meyer-Lindenberg, director of the Central Institute for Mental Health in Mannheim, Germany. And if scientists can work out what aspects of the city are the most stressful, the findings might even help to improve the design of urban areas. "Everyone wants the city to be beautiful but no-one knows what that means," says Meyer-Lindenberg. Wider streets? Taller buildings? More trees? "Architects theorize a lot, but this type of project could deliver a scientific basis for a city code."

RELENTLESS STRESS

Considered from an evolutionary standpoint, the physiological stress response is definitely a good thing: it helps mammals to survive. Any threat, whether from a predator, dwindling food supplies or an aggressive enemy, triggers release of hormones such as cortisol and adrenaline. These hormones raise levels of sugar in the blood and redistribute blood flow to muscles and lungs, so that animals can respond to the threat by running, hunting or fighting.

Problems arise when the stress response doesn't switch off. Stress-hormone levels that stay too high for too long cause high blood pressure and suppress the immune system. And, although the mechanisms are unknown, scientists agree that severe or prolonged stress also raise the risk of psychiatric disease — most brutally in those who have a genetic predisposition, and when the stress occurs while the brain is still developing. In theory, then, the ceaseless challenges of the city could produce this kind of damaging stress. Some fear that they could end up driving an increase in mental illness around the world.

The only signs of an increase, however, come from relatively small, local studies. "It's frustrating — we feel it should be rising," says Ronald Kessler, a mental-health epidemiologist at Harvard Medical School in Boston, Massachusetts. "But globally we have not seen this, and there are also studies which indicate it isn't even rising in cities." However, reliable data on the prevalence of psychiatric disease are hard to find because diagnoses are often imprecise or incompletely recorded. The Camberwell study was influential because, unusually, it captured all those who were diagnosed with a mental disorder, even if they were not admitted to hospitals, and the researchers involved carefully reviewed every case.

Published in 2003, the Camberwell study deeply impressed Meyer-Lindenberg, who was then at the US National Institute for Mental Health in Bethesda, Maryland, researching how genetic risk factors for schizophrenia affect brain function. As a student in Manhattan some years earlier, Meyer-Lindenberg says, "I had been struck by the number of homeless mentally ill people on the streets, and the problems of the city somehow resonated with me". He wondered if city living was somehow making the brain more susceptible to mental-health conditions. When he returned to his native Germany in 2007, he decided to tackle the question directly. But at the time, Meyer-Lindenberg says, "people said the effect would be too subtle to make sense of".

Yet the results of his study, published last year in *Nature* (F. Lederbogen *et al. Nature* **474,** 498–501; 2011), clearly showed that people who grow up in cities process negative emotions such as stress differently from those who move to the city as adults. His team scanned the brains of 55 healthy volunteers as they carried out arithmetic tasks under a constant bombardment of negative social feedback. "We'd always let them know through headphones that we thought they were failing, or at least not doing as well as other subjects we'd had in the scanner," says Meyer-Lindenberg. "In one set of experiments we let them see our impatient faces on computer screens."

T. BATINK

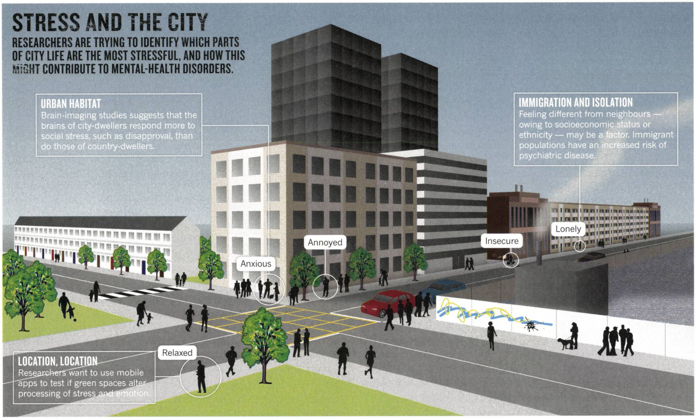
Rush hour: phone apps are helping researchers to survey stress over the day.

This social stress activated two brain areas — but the pattern depended on the volunteers' histories of urban living. The amygdala, which processes emotion, showed much greater activity in people who were currently living in a city. And the cingulate cortex, which helps to regulate the amygdala and processes negative emotion, responded more strongly in those brought up in large cities than in those brought up in the countryside, irrespective of where they lived now. Meyer-Lindenberg thinks that this over-responsiveness to stress could make city-dwellers more prone to psychiatric conditions such as schizophrenia — and his results chime with the idea that stress in childhood or adolescence can have a lasting effect on the brain's development and increase susceptibility to psychiatric disease.

Other scientists are following Meyer-Lindenberg's lead. Daniel Weinberger, director of the Lieber Institute for Brain Development in Baltimore, Maryland (and a self-confessed addict "to the cultural stimulation of the city"), is planning a huge, long-term project to study environmental and genetic risk factors for schizophrenia in China, where urbanization is happening at lightning speed. The proportion of people living in cities there has doubled in the past two decades, to more than half. Together with colleagues at Peking University in Beijing, Weinberger hopes to study thousands of people who moved to Beijing from the countryside before they were 12 years old, after they turned 18 and between the two ages. He will use brain imaging and genetic analyses to try to understand how urban upbringing and genes alter cognition and reasoning, functions that are often disrupted in schizophrenia.

> "THIS TYPE OF PROJECT COULD DELIVER A SCIENTIFIC BASIS FOR A CITY CODE."

Researchers suspect that the stress of city living leads to psychiatric disease mainly in people who are already at risk because of other environmental stresses or because they carry risk genes. One candidate gene, the details of which have not yet been published, has emerged from a large brain-imaging study being conducted by Meyer-Lindenberg in Iceland. He chose more than 500 people who had been identified by the Reykjavik-based company deCODE Genetics as carrying rare mutations that confer a high risk of schizophrenia, and subjected them to functional magnetic resonance imaging tests similar to the ones he used in his city study. "We've already found that people carrying that one particular gene variant activate the cingulate cortex when they process social stress, just like those who were brought up in cities,"



STRESS AND THE CITY

RESEARCHERS ARE TRYING TO IDENTIFY WHICH PARTS OF CITY LIFE ARE THE MOST STRESSFUL, AND HOW THIS MIGHT CONTRIBUTE TO MENTAL-HEALTH DISORDERS.

URBAN HABITAT
Brain-imaging studies suggests that the brains of city-dwellers respond more to social stress, such as disapproval, than do those of country-dwellers.

IMMIGRATION AND ISOLATION
Feeling different from neighbours — owing to socioeconomic status or ethnicity — may be a factor. Immigrant populations have an increased risk of psychiatric disease.

LOCATION, LOCATION
Researchers want to use mobile apps to test if green spaces alter processing of stress and emotion.

he says. He expects to find many more gene candidates through the project, which will run for several more years.

Identifying which parts of a busy city life are the most stressful is another massive challenge (see 'Stress and the city'). The common urban experience of feeling different from your neighbours because of socioeconomic status or ethnicity could be one factor, Meyer-Lindenberg thinks. If so, immigrant groups, who often experience isolation, may be processing stress in a similar way to city-dwellers. He is now testing this hypothesis in the children of immigrants to Germany. (First-generation immigrants are not suitable because the stress of social isolation could be confounded by the stress of moving country.)

Jim van Os, a psychiatrist and epidemiologist at Maastricht University in the Netherlands, is planning a detailed look at city living to identify sources of stress. "It had been slowly dawning on me that as we know that the brain interacts with the environment, nothing about mental health will become clear unless we can look at the environment," he says. Van Os has developed a smart-phone app that allows participants to record their moods, thoughts, location and activities as they go about their daily lives. "This is important because mood and affect change dynamically in the brain, just as blood pressure does," he says.

> **"NOTHING IN MENTAL HEALTH WILL BECOME CLEAR UNLESS WE CAN LOOK AT THE ENVIRONMENT."**

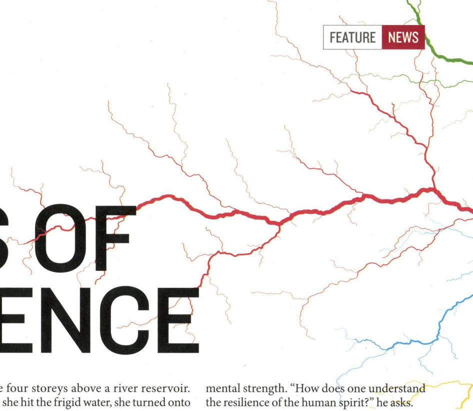
In a €2-million (US$2.6-million) field study, van Os intends to use the app to collect regular information on these factors from 264 people who have started to show psychological disturbance. He will combine these data with brain imaging to test his hypothesis that risk of progressing from disturbance to full-blown psychosis is connected to a distorted ability to learn which aspects of a new environment are rewarding and which are threatening. "For example," he says, "if you move to a new district, you'll need to quickly learn which neighbours you like enough to develop a relationship with, and how to interpret social signals that could be harmful."

Meyer-Lindenberg is planning an even more technologically ambitious project with geoscientists at the nearby University of Heidelberg, who have generated a high-resolution map of their city, and physicists at the Karlsruhe Institute of Technology in Germany, who have developed a mobile device that allows people to be tracked and tested for a week as they walk and work around Heidelberg. The device can recognize when participants reach a specific location — such as a green space or a particularly noisy intersection — and instantly question them about their state of mind or send them a cognitive test to be completed on the spot. The scientists will then ask the participants to come into the lab for brain-imaging studies that examine how they process stress and emotion. By correlating the imaging data with their states of mind at different locations, the team hopes to trace how different aspects of city life affect the brain — whether, for example, strolling through a park really does have a calming influence on the amygdala and cingulate cortex.

It is a high-risk project that has yet to tempt funders. But Meyer-Lindenberg sees the future of the city in it. So too does Annette Rudolph-Cleff, an architect and city planner at the Technical University of Darmstadt in Germany, who contacted Meyer-Lindenberg after reading his *Nature* paper last year and is now advising him on the project. "We know far too little about the city at the moment and we need these new technologies and approaches to help us make decisions about how the city should best be developed," she says.

As well as helping in the design of future cities, such work might also pinpoint the most stressful parts of an existing metropolis — and help to make a case for urban regeneration. Cities are already great economic and cultural incubators; Rudolf-Cleff hopes that the new science of urban stress could also allow them to be turned into cradles of mental health. ■ **SEE EDITORIAL P. 143**

Alison Abbott *is* Nature*'s senior European correspondent.*



THE ROOTS OF RESILIENCE

Most people bounce back from trauma — but some never recover. Scientists are trying to work out what underlies the difference.

BY VIRGINIA HUGHES

On a chilly, January night in 1986, Elizabeth Ebaugh carried a bag of groceries across the quiet car park of a shopping plaza in the suburbs of Washington DC. She got into her car and tossed the bag onto the empty passenger seat. But as she tried to close the door, she found it blocked by a slight, unkempt man with a big knife. He forced her to slide over and took her place behind the wheel.

The man drove aimlessly along country roads, ranting about his girlfriend's infidelity and the time he had spent in jail. Ebaugh, a psychotherapist who was 30 years old at the time, used her training to try to calm the man and negotiate her freedom. But after several hours and a few stops, he took her to a motel, watched a pornographic film and raped her. Then he forced her back into the car.

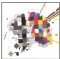
She pleaded with him to let her go, and he said that he would. So when he stopped on a bridge at around 2 a.m. and told her to get out, she thought she was free. Then he motioned for her to jump. "That's the time where my system, I think, just lost it," Ebaugh recalls. Succumbing to the terror and exhaustion of the night, she fainted.

Ebaugh awoke in freefall. The man had thrown her, limp and handcuffed, off the bridge four storeys above a river reservoir. When she hit the frigid water, she turned onto her back and started kicking. "At that point, there was no part of me that thought I wasn't going to make it," she says.

Few people will experience psychological and physical abuse as terrible as the abuse Ebaugh endured that night. But extreme stress is not unusual. In the United States, an estimated 50–60% of people will experience a traumatic event at some point in their lives, whether through military combat, assault, a serious car accident or a natural disaster. Acute stress triggers an intense physiological response and cements an association in the brain's circuits between the event and fear. If this association lingers for more than a month, as it does for about 8% of trauma victims, it is considered to be post-traumatic stress disorder (PTSD). The three main criteria for diagnosis are recurring and frightening memories, avoidance of any potential triggers for such memories and a heightened state of arousal.

Ebaugh experienced these symptoms in the months after her attack and was diagnosed with PTSD. But with the help of friends, psychologists and spiritual practices, she recovered. After about five years, she no longer met the criteria for the disorder. She opened her own private practice, married and had a son.

About two-thirds of people diagnosed with PTSD eventually recover. "The vast majority of people actually do OK in the face of horrendous stresses and traumas," says Robert Ursano, director of the Center for the Study of Traumatic Stress at the Uniformed Services University of the Health Sciences in Bethesda, Maryland. Ursano and other researchers want to know what underlies people's mental strength. "How does one understand the resilience of the human spirit?" he asks.



STRESS AND RESILIENCE
The links between adversity and mental illness. **nature.com/stress**

Since the 1970s, scientists have learned that several psychosocial factors — such as strong social networks, recalling and confronting fears and an optimistic outlook — help people to recover. But today, scientists in the field are searching for the biological factors involved. Some have found specific genetic variants in humans and in animals that influence an individual's odds of developing PTSD. Other groups are investigating how the body and brain change during the recovery process and why psychological interventions do not always work. The hope is that this research might lead to therapies that enhance resilience.

A NATURAL RESPONSE

Although no one can fully understand what was going on in Ebaugh's mind during her attack, scientists have some idea of what was happening to her body. As soon as Ebaugh saw her attacker and his knife, her brain's pituitary gland sent signals to her adrenal glands, atop the kidneys, to start pumping out the stress hormones adrenaline and cortisol. In turn, her pulse quickened, her blood pressure rose and beads of sweat formed on her skin. Her senses sharpened and her neural circuits formed strong memories, so that if she ever encountered this threat in the future, she would remember the fear and flee.

The repercussions were profound. For the first week after the abduction, "I felt like a newborn baby", Ebaugh says, "like I had to be held, or at least be in the presence of somebody". She shivered constantly, was easily startled and felt only fear. She could not go near the grocery store.

Nearly every trauma victim experiences PTSD symptoms to some degree. Many people who are diagnosed with the disorder go on to have severe depression, substance-abuse problems or suicidal thoughts. PTSD can take a horrific toll. Between 2005 and 2009, as a growing number of soldiers faced multiple deployments in Iraq and Afghanistan, suicide rates in the US Army and Marines nearly doubled.

Over the past two decades, researchers have used various kinds of imaging techniques to peer inside the brains of trauma victims. These studies report that in people with PTSD, two areas of the brain that are sensitive to stress shrink: the hippocampus, a deep region in the limbic system important for memory, and the anterior cingulate cortex (ACC), a part of the prefrontal cortex that is involved in reasoning and decision-making. Functional magnetic resonance imaging (fMRI), which tracks blood flow in the brain, has revealed that when people who have PTSD are reminded of the trauma, they tend to have an underactive prefrontal cortex and an overactive amygdala, another limbic brain region, which processes fear and emotion (see 'The signature of stress').

People who experience trauma but do not develop PTSD, on the other hand, show more activity in the prefrontal cortex. In August[1], Kerry Ressler, a neuroscientist at Emory University in Atlanta, Georgia, and his colleagues showed that these resilient individuals have stronger physical connections between the ACC and the hippocampus. This suggests that resilience depends partly on communication between the reasoning circuitry in the cortex and the emotional circuitry of the limbic system. "It's as if [resilient people] can have a very healthy response to negative stimuli," says Dennis Charney, a psychiatrist at the Mount Sinai School of Medicine in New York, who has conducted several brain-imaging studies of rape victims, soldiers and other trauma survivors.

ENVIRONMENTAL PROTECTION

After her abduction, Ebaugh began seeing a psychotherapist and several alternative-medicine practitioners. But more than anything else, she attributes her resilience to being surrounded by caring people — beginning within minutes of her escape.

After Ebaugh crawled up the rocky riverbank, a truck driver picked her up, took her to a nearby convenience store and bought her a cup of hot tea. Police, when they arrived, were sympathetic and patient. The doctor at the hospital, she says, treated her like a daughter. A close friend took her in for a time. And her family offered reassurance and emotional support. "For the first month, I almost had to tell people to stop coming because I was so surrounded by friends and community," she says.

Studies of many kinds of trauma have shown that social support is a strong buffer against PTSD and other psychological problems. James Coan, a psychologist at the University of Virginia in Charlottesville, has done a series of experiments in which women lie in an fMRI scanner and see 'threat cues' on a screen. They are told that between 4 and 10 seconds later, they may receive a small electric shock on the ankle. The cue triggers sensory arousal and activates brain regions associated with fear and anxiety, but when the women hold the hands of their husbands[2] or friends[3], these responses diminish.

Social interactions are complex and involve many brain circuits and chemicals; no one knows exactly why they provide relief. Being touched by someone is thought to stimulate the release of natural opioids, such as endorphins, in the brain. The ACC is packed with opioid receptors, suggesting that touch could influence its response to stress.

Other clues come from the hormone oxytocin, which courses through the brain during social interaction and has been shown to boost trust and reduce anxiety. In one imaging study[4], participants viewed frightening images after receiving nasal sprays of either oxytocin or a placebo. Those who sniffed oxytocin showed reduced activation in the amygdala and weaker connections between the amygdala and the brainstem, which control some stress responses, such as heart rate. The oxytocin surge that comes from being around other people could, like endorphins, help to reduce the stress response.

Past social interactions may also affect how a person responds to trauma. Chronic neglect and abuse unquestionably lead to a host of psychological problems and a greater risk of PTSD. Ressler, however, points to a factor that is well recognized but poorly understood: 'stress inoculation'. Researchers have found that rodents[5] and monkeys[6], at least, are more resilient later in life if they experience isolated stress events, such as a shock or a brief separation from their mothers, early in infancy.

Ebaugh says that early stress — and the confidence she gained in conquering it — helped her to recover from her traumatic abduction. She was born with a condition that made her feet turn inwards. At age ten, she underwent surgery to rebuild her knees followed by a year of intensive rehabilitation. "It wasn't foreign to me to be hurt and have to walk the walk of being strong again," she says. "It's like a muscle, I think, that gets built up."

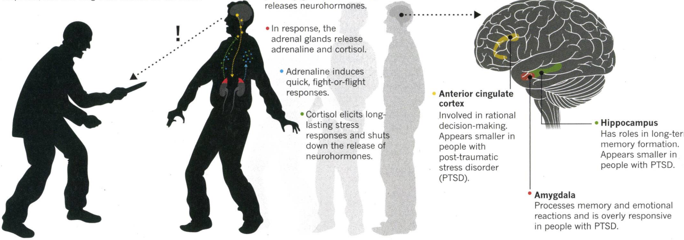
RESILIENT BY NATURE

Although most people, like Ebaugh, recover from trauma, some never do. Some scientists are seeking explanations for such differences in the epigenome, the chemical modifications that help to switch genes on and off (see page 171). Others are looking in the genes themselves. Take, for example, *FKBP5*, a gene involved in hormonal feedback loops in the brain that drive the stress response. In 2008, Ressler and his colleagues showed that in low-income, inner-city residents who had been physically or sexually abused as children, certain variants in *FKBP5* predisposed them to developing PTSD symptoms in adulthood. Other variants offered protection[7].

The most talked-about biological marker of

THE SIGNATURE OF STRESS

A traumatic experience elicits an immediate physiological response, and has long-term effects on the brain.



- During stressful situations, the pituitary gland, in the brain, releases neurohormones.
- In response, the adrenal glands release adrenaline and cortisol.
- Adrenaline induces quick, fight-or-flight responses.
- Cortisol elicits long-lasting stress responses and shuts down the release of neurohormones.

Anterior cingulate cortex
Involved in rational decision-making. Appears smaller in people with post-traumatic stress disorder (PTSD).

Hippocampus
Has roles in long-term memory formation. Appears smaller in people with PTSD.

Amygdala
Processes memory and emotional reactions and is overly responsive in people with PTSD.

E. LARSEN/REDUX

Elizabeth Ebaugh is finally comfortable visiting the bridge from which she was thrown 26 years ago.

resilience is neuropeptide Y (NPY), a hormone released in the brain during stress. Unlike the stress hormones that put the body on high alert in response to trauma, NPY acts at receptors in several parts of the brain — including the amygdala, prefrontal cortex, hippocampus and brainstem — to help shut off the alarm. “In resiliency, these brake systems are turning out to be the most relevant,” says Renu Sah, a neuroscientist at the University of Cincinnati in Ohio.

Interest in NPY and resilience took off in 2000, partly because of a study of healthy US Army soldiers who participated in a survival course designed to simulate the conditions endured by prisoners of war, such as food and sleep deprivation, isolation and intense interrogations[8]. NPY levels went up in the soldiers’ blood within hours of the interrogations. Special Forces soldiers who had trained to be resilient had significantly higher NPY levels than typical soldiers.

Researchers are now conducting animal experiments to study how NPY works. In one experiment, a team at the Indiana University School of Medicine in Indianapolis restrained a rat in a tight-fitting plastic pouch for 30 minutes, then released it into a box with another rat[9]. The restraint made the rat so anxious that it avoided interacting with the other animal for 90 minutes. But when rats were injected with NPY before the treatment, they interacted with cage-mates as if nothing had happened.

The work could lead to treatments. Charney’s group at Mount Sinai is carrying out a phase II clinical trial of an NPY nasal spray for individuals with PTSD. Others are investigating small molecules that can cross the blood–brain barrier and block certain receptors that control NPY release.

CONFLICT RESOLUTION

The US military is leading the hunt for additional biological markers of resilience. Since 2008 — driven in part by soaring suicide rates among soldiers — the US Army has collaborated with the National Institute of Mental Health and several academic institutions on a US$65-million project called Army STARRS (the Study to Assess Risk and Resilience in Servicemembers). The project has many parts, including a retrospective look at de-identified medical and administrative records for 1.6 million soldiers, in search of early warnings of suicide, PTSD and other mental-health problems. STARRS scientists are also collecting data — such as blood samples, medical histories and cognitive testing results — on tens of thousands of current soldiers. The researchers expect to publish their first findings early next year.

The military also funds research into animal models of resilience. Most rodents will quickly learn to associate painful foot shocks with a certain cue, such as a tone or a specific cage. After they have learned the association, the rodents freeze on experiencing the cue, even without the shock. Several years ago, Abraham Palmer, a geneticist now at the University of Chicago in Illinois, made a line of resilient mice by selectively breeding mice that froze for abnormally short periods of time. After about four generations, he had mice that froze for about half the time of typical animals[10]. The effect was not due to a difference in pain sensitivity or general learning ability. This month, Luke Johnson, a neuroscientist at the Uniformed Services University, will present data at the Society for Neuroscience meeting in New Orleans, Louisiana, showing that these mice have uncommonly low activity in the amygdala and hippocampus, consistent with human studies of PTSD resilience. They also have low levels of corticosterone, a stress hormone, in their urine.

“They have a quieter system, even at rest,” says Johnson. “It suggests that there are underlying biological traits that are associated with the capacity of the animal for fear memory.” In future experiments, Johnson plans to use the mice to study NPY and potential new therapies.

Ebaugh, who now specializes in therapy for trauma victims, agrees that drug-based treatments could aid in recovery. But some people may find relief elsewhere. Religious practices — especially those that emphasize altruism, community and having a purpose in life — have been found to help trauma victims to overcome PTSD. Ebaugh says that yoga, meditation, natural remedies and acupuncture worked for her.

Today, she buys groceries at the plaza where she was abducted, and she drives over the bridge she was thrown from as though it were any other road. She says that she has forgiven the man who abducted her. When she reflects on what he did, it’s not with anger, sadness or fear. “It doesn’t feel like it affects my life at all at this point, at least not in a negative way,” she says. “In a positive way, it’s been a huge teacher.” ■

Virginia Hughes *is a freelance science writer in New York city.*

1. Fani, N. *et al. Neuropsychopharmacology* http://dx.doi.org/10.1038/npp.2012.146 (2012).
2. Coan, J. A., Schaefer, H. S. & Davidson, R. J. *Psychol. Sci.* **17,** 1032–1039 (2006).
3. Coan, J. A., Beckes, L. & Allen J. P. *Int. J. Psychophysiol.* (in the press).
4. Kirsch, P. *et al. J. Neurosci.* **25,** 11489–11493 (2005).
5. Maier, S. F. *et al. Dialogues Clin. Neurosci.* **8,** 397–406 (2006).
6. Lyons, D. M. & Parker, K. J. *J. Traum. Stress* **20,** 423–433 (2007).
7. Binder, E. B. *et al. J. Am. Med. Assoc.* **299,** 1291–1305 (2008).
8. Morgan, C. A. III *et al. Biol. Psychiatry* **47,** 902–909 (2000).
9. Sajdyk, T. J. *et al. J. Neurosci.* **28,** 893–903 (2008).
10. Ponder, C. A. *et al. Genes Brain Behav.* **6,** 736–749 (2007).

Flow Cytometry
by
Stratedigm
Cytometry Without Compromise

S1000Ex -
Up to 4 lasers, 14 colors

S1000 -
Up to 4 lasers, 8 colors

SE520 -
2 lasers, 4 colors

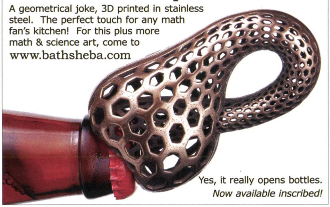
A600 -
High throughput autosampler

Industry leading sensitivity in a compact footprint.

Stratedigm
Cytometry Without Compromise

www.stratedigm.com
408.512.3910 info@stratedigm.com

The Klein Bottle Opener
A geometrical joke, 3D printed in stainless steel. The perfect touch for any math fan's kitchen! For this plus more math & science art, come to
www.bathsheba.com



Yes, it really opens bottles.
Now available inscribed!

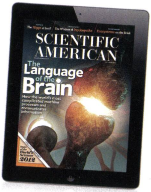
CUSTOM MONOCLONAL ANTIBODIES
Rat mAb ◦ Mouse mAb ◦ Ascites

CUSTOM PROTEIN EXPRESSION
HEK293 ◦ Baculovirus ◦ *E. coli*

Superior Quality • Unrivaled Expertise
www.promab.com
1.866.339.0871 | info@promab.com
ProMab
Biotechnologies, Inc.

PROMOTION

Expand your mind.



Get closer to the leading edge of science and technology: experience the SCIENTIFIC AMERICAN Tablet Edition - now available for iPad®.

Available on the App Store

TO RENEW
SCIENTIFICAMERICAN.COM/UPGRADE

TO SUBSCRIBE
SCIENTIFICAMERICAN.COM/TABLET

Apple and iPad are trademarks of Apple Inc., registered in the U.S. and other countries. App Store is a service mark of Apple Inc.

COMMENT

STRESS Epigenetics may set resilient and vulnerable people apart **p.171**

HEALTH A call to halt the medicalization of sleep **p.173**

FILM Award-winning movie explores the dark past of US psychiatric care **p.174**

JOURNALS Are cash incentives inflating publications and citations in China? **p.176**

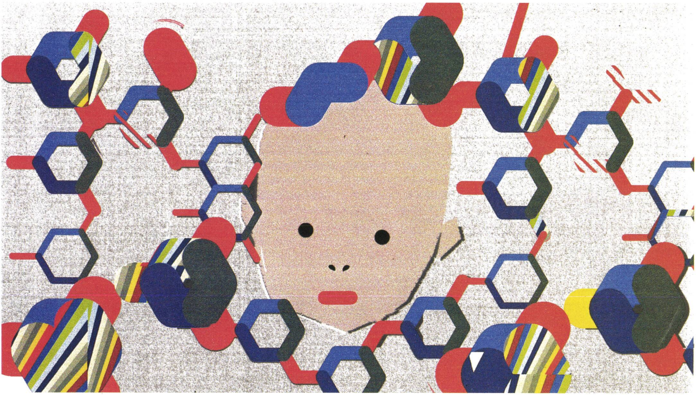


ILLUSTRATION BY PADDY MILLS

Too toxic to ignore

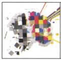
A stark warning about the societal costs of stress comes from links between shortened telomeres, chronic stress and disease, say **Elizabeth H. Blackburn** and **Elissa S. Epel**.

In the 2006 film *The Holiday*, the actress Cameron Diaz, playing a woman whose life is spinning out of control, exclaims: "Severe stress ... causes the DNA in our cells to shrink until they can no longer replicate. So when we're stressed we look haggard."

Hollywood got that science right. The DNA to which Diaz's character alludes is the segment that makes up telomeres, structures that cap and protect the ends of chromosomes. She was referring to our 2004 publication[1] — the first to link chronic psychological stress to compromised telomere maintenance.

Since that paper, researchers have consistently found that various types of chronic stress are linked to — and probably cause — shorter telomeres. Meanwhile, both telomere shortness and stress have independently been associated with several common conditions, such as cardiovascular disease and diabetes.

These associations are so widespread and consistent that even without a detailed understanding of the biochemical pathways involved, the message is clear. Failure to alleviate severe stress caused by prolonged threats such as war, financial hardship, abuse and emotional neglect, particularly in children, will result in exponentially higher costs further down the line — personal, economic and otherwise.



STRESS AND RESILIENCE
The links between adversity and mental illness. **nature.com/stress**

SHORT ODDS

Human telomeres are several kilobases of repeated sequences of DNA bound by specialized protective proteins. A peculiarity of the DNA-replication mechanism causes telomeres to shorten as cells divide. Sometimes the enzyme telomerase can replenish the lost DNA, but as we age, our telomeres dwindle. If they get too short, through ageing or because telomere maintenance goes awry, cells can stop dividing. Such cells also become malfunctional. For instance, they can start secreting factors that cause inflammation or trigger the development of tumours. ▶

In 2004, we compared telomere lengths in the white blood cells of mothers of chronically ill children to those in mothers of healthy children[1]. The longer a woman had spent being the main carer of her ill child (the children's conditions ranged from gut disorders to autism), the shorter were her telomeres. Moreover, in both groups, the more severe her psychological stress — as judged by her answers to standardized questions about, for instance, how in control she felt over her life — the shorter were her telomeres. The extra telomere shortening in the 'most stressed' mothers (compared with that in the 'least stressed' mothers) was equivalent to that caused by at least a decade of ageing.

> *"Telomeres powerfully quantify life's insults."*

This relationship between stress and telomere length keeps showing up: from studies of kindergarten-aged children to adults as old as 80; from small clinical samples of less than 100 people to large population-based samples of thousands[2,3].

Under stress, the body ramps up its production of certain hormones, such as cortisol, and other biochemical factors. These compounds help to mediate an appropriate response to short-term stress. But when overproduced for months or years, they can alter gene expression, probably with deleterious effects (see page 171). In the laboratory, the same factors can shorten telomeres — in the case of cortisol, by reducing the activity of telomerase. It is likely that the pathways that mediate alterations to gene expression interact with those affecting telomere maintenance, although this has yet to be explored.

Although many studies have unearthed (and continue to uncover) associations between stress and eroded telomeres, others have forged links between telomere shortness and common disorders. Rare mutations of genes encoding components of telomerase cause telomeres to be too short. This results in immune-system disorders such as aplastic anaemia, and other conditions including pulmonary fibrosis, diabetes, some cardiovascular diseases and certain cancers. Remarkably, many of these inherited diseases, recently named 'telomere syndromes'[4], are those that are commonly associated with ageing.

Telomere shortness can even predict people's statistical risk of developing certain conditions. For instance, for 10 years after their telomeres had been measured, men and women with shorter telomeres were three times more likely than those with longer ones to develop certain types of cancers such as pancreatic cancer[5]. If elderly, people with shorter telomeres were also 50% more likely than those with longer telomeres to develop dementia and 50% more likely to die from any particular cause[6].

Evidence that lifestyle, well being and other environmental factors can contribute significantly to disease has been accumulating for decades. But we now have three pairwise links involving three factors: stress with telomere shortness; stress with disease risks; and telomere shortness with risks for these diseases. It is hard to avoid the inference that at least one of the ways stress causes chronic diseases is by shortening telomeres.

TIME TO ACT

How can policy-makers act now on the wealth of evidence that is accumulating from studies of telomeres?

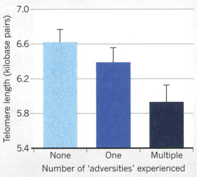
Dissecting the cellular and physiological pathways linking unusually short telomeres to stress and disease could prove important for finding possible targets for treatments. But policy-makers don't need to wait for all the mechanistic details to be filled in (which could take decades) — especially given that what happens in the laboratory often fails to reflect what is going on in the body as a whole.

A striking message from the telomere story is the importance of considering environmental as well as genetic effects in addressing disease. From 2008 to 2011, the US National Institutes of Health (NIH) allocated a total of US$29 million to genetic research. By comparison, over the same period, it gave only $14 million to all behavioural- and social-sciences research. This is despite several analyses attributing roughly 50% of the variance in early mortality in the United States to largely modifiable behaviours such as overeating, alcohol abuse and smoking (which are also stress-related in part)[7].

Using current technologies, people's gene variations are easier to track down than the plethora of influences coming from outside the body, let alone how people behave, think and feel. Yet capturing a person's vulnerability to disease will require understanding all these inputs. Telomere length provides an especially good window into global physiological status and is a bellwether of disease. (Interestingly, it is often more affected by a person's life experience than by which variants of telomere-maintenance genes they carry.)

TELOMERES TELL

They are shorter in adults who experienced more trauma as children.



SOURCE: J. K. KIELCOLT-GLASER *ET AL. PSYCHOSOM. MED.* **73,** 16–22 (2011).

Telomere research also points to some practical ways to improve health. Mice that are genetically deprived of telomerase quickly become wizened and grey but such changes, normally associated with ageing, can be at least partly reversed by restoring telomerase activity[8]. Designing drugs to boost telomerase in humans without inducing unwanted side effects is a formidable challenge. More feasible approaches to alleviating telomere shortening could involve mitigating the conditions that lead to chronic stress and helping people to change certain behaviours.

Encouragingly for this latter approach, some pilot studies suggest that just three months of stress-reduction interventions, often along with increased physical activity and dietary changes, may slow or even reverse telomere attrition by increasing telomerase activity. An intriguing challenge is to see whether the idea of telomeres eroding (which many have told us conjures up a striking image of declining health) could motivate people to change their behaviour.

START YOUNG

Individual efforts apart, perhaps the strongest message to come from the work on telomeres is that to pre-empt many common diseases, especially those that are becoming increasingly prevalent in the world's ageing population, governments and other policy-makers need to prioritize what we call 'societal stress reduction'. Meditation retreats or yoga classes may help those who can afford the time and expense. But we are talking about broad socioeconomic policies to buffer the chronic stressors faced by so many.

Together, the results from studies of telomeres, stress and disease reinforce a 2008 recommendation from the World Health Organization (WHO) for reducing health inequities. The WHO proposes that governments and global health organizations focus on improving education and the conditions surrounding early child development[9].

Several studies indicate that stress begins eroding telomeres in infancy, and maybe even before children are born. For example, studies have found that the more violence children had experienced, or the longer they had spent in an orphanage, the shorter were their telomeres[2]. Even young healthy adults whose mothers had experienced severe stress while pregnant (for instance, because of a close family member dying) had shorter telomeres than people whose mothers had relatively stress-free pregnancies[10].

What is more, the effects of stress early in life reverberate into adulthood. In more than 4,000 middle-aged UK women, for example, the more categories of adversity each woman had experienced as a child (such as physical abuse or parental divorce, unemployment or drug use), the shorter were her telomeres[3].

This has also been found in the elderly (see 'Telomeres tell'). Even experiencing fewer years of school education in early life is associated with having shorter telomeres in middle and old age.

Delaying actions that mitigate diseases such as diabetes until adulthood will only exacerbate personal and societal costs. One example of a proactive step is the US health programme Medicaid's Strong Start initiative, which aims to enhance pregnancy-related care. Improving the education and health of women of child-bearing age in general could be a highly effective way to prevent poor health filtering down through generations.

To suggest that people's quality of life matters or that societies and governments should be allocating more resources to mothers and children is hardly new or controversial. What is new is the wealth of evidence demonstrating that telomeres powerfully quantify life's insults. They are shorter in people who were exposed to adversity as children, and shorter still for each year a person spends depressed, caring for a sick child, being abused and so on.

Telomeres send one more signal — from the tips of our chromosomes — that unmanageable social and psychological stress, especially during early life, is as insidious as smoking or too much fast food. ■ SEE COMMENT P.171

Elizabeth H. Blackburn *is the 2009 Nobel Laureate in Physiology or Medicine and professor of biology and physiology in the Department of Biochemistry and Biophysics.* **Elissa S. Epel** *is associate professor in the Department of Psychiatry, University of California, San Francisco, California 94143, USA, and is part of several National Institute on Aging initiatives on stress research.*
e-mail: elizabeth.blackburn@ucsf.edu

1. Epel, E. S. *et al. Proc. Natl Acad. Sci. USA* **101,** 17312–17315 (2004).
2. Shalev, I. *et al. Mol. Psychiatry* http://dx.doi.org/10.1038/mp.2012.32 (2012).
3. Surtees, P. G. *et al. J. Gerontol. A Biol. Sci. Med. Sci.* **66,** 1152–1162 (2011).
4. Armanios, M. & Blackburn, E. H. *Nature Rev. Genet.* **13,** 693–704 (2012).
5. Willeit, P., Willeit, J., Kloss-Brandstätter, A., Kronenberg, F. & Kiechl, S. *J. Am. Med. Assoc.* **306,** 42–44 (2011).
6. Honig, L. S., Kang, M. S., Schupf, N., Lee, J. H. & Mayeux, R. *Arch. Neurol.* http://dx.doi.org/10.1001/archneurol.2012.1541 (2012).
7. Mokdad, A. H., Marks, J. S., Stroup, D. F. & Gerberding, J. L. *J. Am. Med. Assoc.* **291,** 1238–1245 (2004).
8. Jaskelioff, M. *et al. Nature* **469,** 102–106 (2011).
9. Commission on Social Determinants of Health. *Closing the Gap in a Generation* (World Health Organization, 2008).
10. Entringer, S. *et al. Proc. Natl Acad. Sci. USA* **108,** E513–E518 (2011).

Competing financial interests declared; see go.nature.com/iznghy for details.

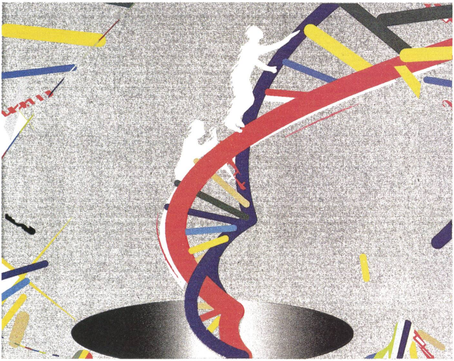


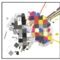
ILLUSTRATION BY PADDY MILLS

Stress makes its molecular mark

Trauma affects people differently. Epigenetics may be partly to blame, says **Eric J. Nestler**.

Some people exposed to severe stress, such as that caused by prolonged economic hardship or sexual or physical abuse, go on to develop devastating psychological or other health problems. Others are more resilient. If one identical twin shows symptoms of stress-related depression, for instance, the other will also be depressed in only around 40% of cases. I believe that epigenetic mechanisms help to explain why[1,2]. These are experience-dependent molecular alterations to DNA or to proteins that alter how genes behave without changing the information they contain.

Recent studies suggest that epigenetic mechanisms shape short-term (lasting hours) and long-term (lasting months, years or even a lifetime) responses to stress. Some studies even hint that epigenetic changes could affect the next generation. A serious effort to both map and substantiate associations between behavioural responses and epigenetic alterations — although costly and challenging — would almost certainly flag up possibilities for treatments that either reverse the effects of stress or enhance a person's ability to cope.



STRESS AND RESILIENCE
The links between adversity and mental illness. nature.com/stress

AGGRESSIVE MICE

When a person is stressed, gene expression in parts of the brain may be up- or down-regulated. This can occur through chemical modifications to DNA, to regulatory proteins in the nuclei of brain cells or to histones (proteins that package and order DNA). Many stress-induced changes are adaptive, but some seem to be damaging.

In my laboratory, we have stressed mice by repeatedly exposing them to more aggressive mice[1] (see 'A switch for resilience'). After ten days of this treatment, the stressed mice begin to avoid other mice, show less interest in things that normally excite them (such as sweets and sex), become less adventurous and even grow obese (they take less pleasure in eating but eat more). ▶

Many of these symptoms can persist for months and are treatable with standard antidepressant medications. We have also found that mice given cocaine the week before being exposed to an aggressive mouse have more extensive epigenetic modifications, which induce more stress-related symptoms[3].

Of the hundreds of mice studied in my lab, roughly one-third become less adventurous when stressed but have no other symptoms. By looking at differences in gene expression and structural organization of DNA between these 'resilient' mice and more susceptible mice, we have linked distinct behavioural responses to specific molecular alterations — all in regions of the brain important in reward recognition[3–6]. These alterations include differences in DNA methylation, patterns of attachment of acetyl or methyl groups to histones and activity of various transcription factors. They last for days or, in some cases, for several weeks.

We can make susceptible mice resilient by blocking or inducing epigenetic modifications to certain genes or by altering the expression patterns of those genes to mimic the epigenetic tweaks. Likewise, epigenetic modifications and gene expression can be altered in resilient mice to make them more susceptible.

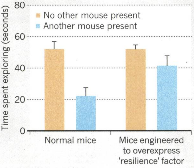
Other groups have found similar epigenetic alterations that last a lifetime. For instance, rat pups that are rarely licked and groomed by their mothers are more susceptible to stress later in life[7] than are pups with more diligent carers. They are less adventurous than better-cared-for offspring and put up less of a fight in unpleasant situations (such as being placed in a beaker of water). Moreover, the females are less nurturing towards their own offspring. Epigenetic modifications seem to occur at several genes in the hippocampus in response to how much grooming young rats receive, and these alterations persist into adulthood[7].

These findings are likely to hold up in humans. For example, researchers have found that the genes identified in the rat-grooming studies were more methylated in the hippocampi of suicide victims who had experienced trauma as children than in the those of people who had died from suicide or natural causes and whose childhoods were normal. Likewise, our findings in mice given cocaine mirror epidemiological studies from the past few decades that have linked drug abuse, obesity and conditions such as multiple sclerosis, diabetes and heart disease to increased susceptibility to stress in humans.

More controversial is whether animals inherit epigenetic vulnerability to stress. According to this notion, epigenetic modifications in sperm or eggs drive aberrant patterns of gene expression in the next generation[8]. Several groups have reported that male mice exposed to stress — by being removed from their mothers as pups or exposed to more aggressive mice as adults, for example — produce offspring that are more vulnerable to stress[9,10].

A SWITCH FOR RESILIENCE

Mice become shy of other mice after repeated exposure to aggressive peers. Mimicking certain epigenetic tweaks makes them bolder.



SOURCE: V. VIALOU *ET AL. NATURE NEUROSCI.* **13**, 745–752 (2010).

A mechanism is still elusive. Exposure to stress could somehow corrupt the male mouse's behaviour or affect some signalling molecule in his semen such that his partner alters her care for their young. Another possibility is that stress-linked epigenetic 'marks' in the sperm affect the development of offspring[9,10]. No causal evidence yet links epigenetic changes in sperm to altered behaviour in offspring.

MAPPING MARKS

Epigenetics is in vogue: over the past five years, researchers have proposed epigenetic explanations for all sorts of phenomena, from language acquisition to obesity, without clear proof. At a meeting I attended two years ago, people proposed that the spread of Christianity in the early centuries AD was partly due to epigenetic mechanisms. Furthermore, researchers too often identify correlations between behaviour and molecular alterations in cells without also establishing a causal link. Some biologists are, rightfully, wary.

Yet the results I have described show how important epigenetic mechanisms are likely to be in understanding the effects of stress and in discovering ways to manage it.

It is time for researchers to begin the difficult business of substantiating associations. Cost is one challenge to pinpointing the genes and biochemical pathways involved in epigenetically mediated responses to stress. Hundreds of known types of modifications probably act in complex combinations. Mapping each mark at specific points in development in a rat or mouse, and doing so in each brain region or peripheral tissue, would cost tens of thousands of dollars. Defining the alterations in the many cell types in a given brain region[2] pushes the cost several-fold higher. And the expense for doing so in humans is orders of magnitude higher still. Human genetic diversity means that researchers will probably need to study hundreds or thousands of people to obtain a meaningful picture. A related challenge is that of obtaining enough computational power to analyse the hundreds of terabytes of sequencing data that would be produced — although recent advances in bioinformatics are beginning to help.

At present, researchers study epigenetic alterations by up- or downregulating enzymes such as histone methyltransferases. But such enzymes can influence thousands of genes. Tools that allow researchers to target a specific type of epigenetic modification to a single gene in a particular cell type *in vivo* will usher in a far more compelling phase of research.

Work in simpler organisms, such as the roundworm *Caenorhabditis elegans*, is providing clues about the range of epigenetic modifications that can occur in sperm or egg cells. Of course, experiments in mammals will be needed to establish whether epigenetic transmission of information can occur across generations to a meaningful extent.

Thirty per cent of lost productivity worldwide is caused by psychiatric conditions such as depression, anxiety and schizophrenia — all of which are exacerbated by chronic stress. In the developed world, that proportion is 40%. The toxic levels of stress people are now exposed to — in part thanks to the increases in productivity, lifespan and competitiveness that come with a wealthier, healthier globalized economy — are here to stay. A serious endeavour to understand why people respond to stressful experiences so differently — challenging as it would be — is easily justified. ■ SEE COMMENT P.169

Eric J. Nestler *is Nash Family Professor at the Fishberg Department of Neuroscience and the Friedman Brain Institute, Mount Sinai School of Medicine, New York, New York 10029, USA.*
e-mail: eric.nestler@mssm.edu

1. Sun, H., Kennedy, P. J. & Nestler, E. J. *Neuropsychopharmacology* advance online publication, http://dx.doi.org/10.1038/npp.2012.73 (13 June 2012).
2. Peter, C. J. & Akbarian, S. *Trends Mol. Med.* **17,** 372–379 (2011).
3. Covington, H. E. III *et al. Neuron* **71,** 656–670 (2011).
4. LaPlant, Q. *et al. Nature Neurosci.* **13,** 1137–1143 (2010).
5. Wilkinson, M. B. *et al. J. Neurosci.* **29,** 7820–7832 (2009).
6. Wilkinson, M. B. *et al. J. Neurosci.* **31,** 9084–9092 (2011).
7. Caldji, C., Hellstrom, I. C., Zhang, T.-Y., Diorio, J. & Meaney, M. J. *FEBS Lett.* **585,** 2049–2058 (2011).
8. Dunn, G. A., Morgan, C. P. & Bale, T. L. *Horm. Behav.* **59,** 290–295 (2011).
9. Franklin, T. B. *et al. Biol. Psychiatry* **68,** 408–415 (2010).
10. Dietz, D. M. *et al. Biol. Psychiatry* **70,** 408–414 (2011).

H. KORVOLA/GETTY

Some people need more rest than the prescribed eight hours a night — or need it at different times.

SLEEP SCIENCE

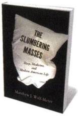
Broken dreams

Meredith Wadman lifts the blanket on the creeping medicalization of sleep in the United States.

Millions of people in the United States struggle to achieve that great American dream, a 'good night's sleep'. So says Matthew Wolf-Meyer in his ambitious *The Slumbering Masses*. He contends that capitalist necessity defines sleep in the country today, shoehorning sleepers into a societally convenient but physiologically arbitrary eight-hour night.

Those who can't manage the prescribed amount of slumber at the prescribed time are often labelled disordered sleepers. Wolf-Meyer's message is that society should bend to accommodate, even celebrate, diversity in sleeping behaviour, rather than branding nonconformism pathological.

For now, night owls and others who fail to adapt to the eight-hour, nocturnal norm — whether owing to disease or a particular hard-wired biology — are a boon to the pharmaceutical industry. Wolf-Meyer, an anthropologist, calls for a shift towards more flexible organization of workdays, school and social lives, and away from the assumption of monolithic "slumbering masses". Otherwise, he warns, "Americans may be doomed to a future of proliferating sleep disorders, amphetamine breakfasts, and sedatives for dinner."



The Slumbering Masses: Sleep, Medicine and Modern American Life
MATTHEW J. WOLF-MEYER
Univ. Minnesota Press: 2012. 312 pp. $24.95, £18.50

Wolf-Meyer lays much responsibility for the medicalization of sleep at the feet of a US sleep-medicine establishment that has grown up since the 1950s. Its roots, however, emerge in his fascinating history of the Protestant origins of sleep in the United States. The influential Puritan minister Cotton Mather argued in the late seventeenth century that those with a proclivity for the luxuries of slumber were failing in their earthly and God-given duty to be productive. The bed, Mather opined, is one of just a few places where "the Devil has laid out most fatal snares". A generation later, Benjamin Franklin turned the same message positive with his still-famous dictum, "Early to bed, early to rise, makes a man healthy, wealthy and wise."

This morality handily converged with the twentieth-century idea of 'normal sleep': a consolidated nocturnal experience programmed by biology, but potentially moulded by behaviour modification and, of course, drugs. The work of sleep-science pioneer Nathaniel Kleitman, who in 1938 descended into Mammoth Cave in Kentucky to try to realign his biology to a 28-hour day, thrust an awareness of sleep research into the public eye. It also reinforced the idea that sleep can be controlled by willpower — a concept further cemented by Kleitman's student, William Dement.

Dement founded the first sleep laboratory, at Stanford University in California, in the early 1970s. He understood healthy sleep "as resting upon a biological foundation of eight quiet, motionless and consolidated hours of sleep … between sunset and sunrise", writes Wolf-Meyer, arguing that Dement's *The Promise of Sleep* (co-written by Christopher Vaughan; Delacorte, 1999) "promotes a model of nature and human biology from which all variations are disorders". Work such as Dement's has, in Wolf-Meyer's view, been one factor in helping to pave the way for the modern pharmaceutical industry to find a broad new market among disordered US sleepers. Many of these sleepers are given the catch-all diagnosis of 'excessive daytime sleepiness', the use of which in the medical literature has exploded in the past ten years.

Those looking for an exhaustive pharmacopeia of the sleep-medicine industry will not find it here. Wolf-Meyer notes that US ▶

STRESS AND RESILIENCE
The links between adversity and mental illness. nature.com/stress

drug-maker Sepracor spent "hundreds of millions of dollars" on the US launch of the sedative Lunesta (eszopiclone). But he offers only one example to support his contention that companies are widening their net: pharmaceutical firm Cephalon's successful effort to expand the user base for the stimulant Provigil (modafinil) from people with narcolepsy to those with sleep apnoea and 'shift-work sleep disorder'. Neither has Wolf-Meyer tracked down the associated and — he implies — growing revenue numbers for leading stimulants and sedatives. Such figures would buttress his claim that in "a few short years, excessive sleepiness may successfully become the new erectile dysfunction".

Wolf-Meyer might have delved deeper if his scope had been narrower. As it is, he covers everything from the socialization of children to sleep at the appropriate hour (think of Margaret Wise Brown's bedtime classic *Goodnight, Moon*) to the use and abuse of caffeine and other stimulants, and the plight of workers in Indian call centres, forced to synchronize their hours with US time. Yet in other ways, the broad treatment pays dividends. It is illuminating, for instance, when Wolf-Meyer takes us inside a Minnesota sleep clinic where medics are struggling to help a sleepless young girl already prescribed five drugs.

Implicit in Wolf-Meyer's analysis is that the medicalization of sleep is a profit-driven pursuit. No doubt economics did indeed have a big role in the rise of sleep medicine, but the author finds little room for the real benefits of some treatments. Can it be bad when medication prevents a narcoleptic from falling asleep at the wheel? And I, as a sleep apnoeic with daily work and family responsibilities, am indebted to the machine that keeps my airway open at night and prevents what would otherwise be literally hundreds of sub-conscious wakenings.

The Slumbering Masses suffers in many places from jargon. Sentences such as, "Normative desire facilitates the functioning of everyday spatiotemporal hegemony and is in turn formed through that very same hegemony" made me long for a lay-friendly translation.

But there are passages of telling clarity. Wolf-Meyer tells the story of Betsy, a woman in her 50s who fought insomnia for decades. She tried, she confesses, "lots and lots of drugs. Everything from benzo[diazepines] to Xanax, antidepressants, and all the tricyclics [as well as] muscle relaxers mixed with other drugs. And they're effective for a while, and then they all wear off." ■

Meredith Wadman *is a correspondent for* Nature *based in Washington DC.*

In the early twentieth century, unpaid patient labour kept Kings Park hospital self-sufficient.

PHOTO: KINGS PARK HERITAGE MUSEUM/DIGITAL RESTORATION: P. A. BLACKSBERG

PSYCHIATRY

The dispossessed

Amy Maxmen views a prizewinning film that shines a light into the dark corners of US psychiatric care.

Lucy Winer checked into Kings Park psychiatric hospital on Long Island, New York, after overdosing on sleeping pills and slitting her wrists. It was 1967; she was 17. Stern nurses dressed her in a hospital gown and escorted her into a room where identically dressed women slept on the floor or leaned lifelessly against walls. The women were literally floored by antipsychotic medications that, as Winer was to find in her six months in the hospital, felt like an iron suit.

"We had been thrown away, stripped, locked up. We were disposable," says Winer, in the documentary *Kings Park*. Winer directed and co-produced the film 30 years after her stay at the hospital, now long abandoned. *Kings Park* tells a tale of mental health care that must be told, she says. The psychiatrists who are now showing it at meetings and workshops around the United States agree: last month, the New York Association of Psychiatric Rehabilitation Services presented Winer with the 2012 Public Education/Media award. *Kings Park* touches a nerve.

The psychiatrists' goal is to inspire progress by conveying to mental health-care providers how it felt to be on the receiving end of deficiencies in state mental institutions. Although the drugs administered to people with serious mental illness are arguably less dangerous now than they were in the 1960s and therapy is widely accepted, not everyone who needs these advances receives them. In the United States, more than 10% of people with serious mental illness are now homeless, or in prison (L. Davis *et al. Curr. Psychiatry Rep.* **14,** 259–269; 2012).

Kings Park: Stories From an American Mental Institution
DIRECTED BY LUCY WINER
Wildlight Productions: 2012.
kingsparkmovie.com

Winer's intent, too, is to shine a light on the dehumanized patients — and overwhelmed doctors — challenged by inadequate resources for mental-health treatment. She spent 11 years creating the film to explore her terrifying personal experience with mental illness as a teenager. As she turns her gaze outwards in interviews with psychiatrists, attendants and other former patients, *Kings Park* also becomes a history of US psychiatry told from multiple perspectives. The journey ends in the present, at the assisted-living centres and prisons where many former residents of psychiatric hospitals live. Most large state-run mental hospitals in the United States shut their doors over the course of four decades: between 1955 and 2003, the number of inpatients in them dropped by more than 90%.

STRESS AND RESILIENCE
The links between adversity and mental illness. nature.com/stress

Kings Park was comparable to other US state hospitals, but larger and more venerable than many. It opened in 1885 as an asylum for people with mental illness in New York City. In the 1950s, it housed roughly 9,000 patients at any one time. In the same decade, Thorazine (chlorpromazine), an antipsychotic drug prescribed for schizophrenia and other psychiatric disorders, came to Kings Park. It replaced cruder means of quieting patients, such as lobotomy and electric-shock therapy, but caused slurred speech, the writhing and shaking of tardive dyskinesia and other distressing side effects, as doctors experimented with various formulations.

Understaffing occasionally led to patients being harmed when attendants rushed to quiet them. Hannelore Lehnhoff, a psychiatrist at Kings Park from 1960 to 1985, is noticeably distraught as she recalls an extreme example in which a patient seemed to have been suffocated with a pillow by nurses. Winer captures the frustration that psychiatrists felt with too little time to address the needs of so many patients, and a lack of tools to tend to those they saw. The sombre tone lifts momentarily with the news of the Community Mental Health Act signed by President John Kennedy in 1963, which outlined how small, assisted-living situations could provide an alternative to massive state-run institutions. Funding to Kings Park fell soon after the act was passed, and the hospital began to downsize. It shut in 1996.

At the assisted-living facilities visited in *Kings Park*, former patients cook, give tours to prospective residents and their families, and discuss in peer-support groups how much better it is to be needed than to be abandoned.

But such facilities were and remain rare. *Kings Park* shows how federal and state budgets were scaled back as state hospitals closed, and many patients ended up on the streets. A trembling former Kings Park resident with schizophrenia, who is now in prison, speaks to camera about how he was arrested for sleeping outside a church. In scenes at Suffolk County Correctional Facility in New York, we hear that one-fifth of the inmates have been diagnosed as mentally ill. Once more, their clothes have been swapped for uniforms.

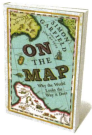
After a screening at the American Psychiatric Association's annual meeting in May, the discussion lasted for more than an hour. For Evelyn Bromet, a professor of psychiatry at Stony Brook School of Medicine in New York, *Kings Park* offers a provocative way to teach history to avoid repeating it. "There are mental-health researchers who have no appreciation for what state hospitals were like," says Bromet. "Winer is telling the story of an enormous group of people who are forgotten." ■

Amy Maxmen *is a freelance journalist in Brooklyn, New York.*
e-mail: amaxmen@gmail.com

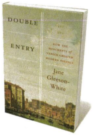
Books in brief

Bad Pharma
Ben Goldacre
Bestselling author of Bad Science
How drug companies mislead doctors and harm patients

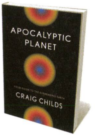
Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients
Ben Goldacre FOURTH ESTATE *448 pp. £13.99 (2012)*
Psychiatrist and firebrand Ben Goldacre puts big pharma in the dock. Skewering an industry riddled with 'side effects' — from suppressed trial results to diseases invented for profit — and backed by poor regulation and aggressive marketing, Goldacre also offers pragmatic solutions. Further exposure of a seam mined by, among others, Marcia Angell, former editor-in-chief of *The New England Journal of Medicine* (*The Truth about the Drug Companies*; Random House, 2004). Hear an interview with Goldacre at go.nature.com/huccmd.



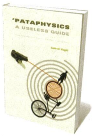
On The Map: Why the World Looks the Way it Does
Simon Garfield PROFILE BOOKS *468 pp. £16.99 (2012)*
The sixteenth-century Mercator world projection has aged well — Google Maps uses it. Cartography, asserts Simon Garfield, evolves endlessly but is rooted in enduring needs: to discover frontiers, plot progress, keep our bearings. The myriad maps he shows are doorways into key moments in cartographic history, from Venetian monk Fra Mauro's 1459 world map, the last hurrah of the medieval planisphere, to Gordon Home's ruler-like rendering of Roald Amundsen's 1911 route to the South Pole. Disease mapping, brain mapping, gaming — charting the world has morphed into mapping worlds within worlds.



Double Entry: How the Merchants of Venice Created Modern Finance
Jane Gleeson-White ALLEN & UNWIN *304 pp. £12.99 (2012)*
Intrigued by the economic engine driving Renaissance art, Jane Gleeson-White crafted this gem of a history. It hinges on monk Luca Pacioli, who published the first treatise on double-entry bookkeeping — a mercantile system underpinning today's global economy. There are intriguing circularities: Pacioli probably learned mathematics from artist Piero della Francesca, then helped Leonardo da Vinci with linear perspective. Gleeson-White traces the system's impact through Keynesian successes and today's high-finance excesses.



Apocalyptic Planet: Field Guide to the Everending Earth
Craig Childs PANTHEON BOOKS *368 pp. £17.46 (2012)*
Mass extinctions, eras of "equatorial seas bobbing with slush", wall-to-wall desert — Earth's cycles of destruction, science writer Craig Childs reminds us, are dramatic, relentless and constant. Which end of the world will be next? Childs travelled to nine apocalyptic places for a taste of possible future cataclysms. Each snapshot is research-rich: we get the swirling of warm air cells towards the poles on a trek in North America's Sonoran Desert, and the "genetic exhaustion" of Iowa on a gruelling hike through vast cornfields. Science shot through with real lyricism.



'Pataphysics: A Useless Guide
Andrew Hugill MIT PRESS *296 pp. £17.95 (2012)*
For a science that doesn't exist, 'pataphysics has popped up with happy regularity for decades — ever since playwright and exquisite jokester Alfred Jarry conceived of the proto-Dadaist 'science of imaginary solutions'. Andrew Hugill, a professor of digital humanities, has spent 25 years charting 'pataphysics in science and the arts. He teases out its influence on everything from the philosophy of Jean Baudrillard to the novels of Philip K. Dick — and suggests that Jarry may have pre-empted theories of antimatter.

Correspondence

Drop re-review for big grant holders

The fundamental concern with a second layer of review for holders of more than US$1 million in grants from the US National Institutes of Health is that it takes us further from a meritocracy in which the best and highest-impact science is supported (*Nature* **489,** 203; 2012).

Elite grant recipients have gone through reviewers' scrutiny and proved their productivity many times. Reviewers and grant administrators already check for potential funding overlap, one of the issues the reforms are designed to address. Peer review may be imperfect, but it is done by the scientific leaders in the relevant field, and there is no reason to believe that the judgement of an advisory council is superior.

The reluctance of most funding agencies to pay for research infrastructure necessitates either leveraging economies of scale from multiple grants or belonging to an academic institution that is wealthy enough to share the costs — a position not all researchers enjoy. The new policy would mainly affect the most productive principal investigators with multiple R01-type (individual project application) grants, widely considered to yield the most innovative research.

Taxpayers have the right to make the scientific community accountable for improving health, but this extra scrutiny may compromise attainment of that goal.

Hemant K. Roy *Northwestern University, Evanston, Illinois, USA. h-roy@northwestern.edu*

EU scientific visa is a success in Ireland

From experience in Ireland, I contend that Europe's scientific visa and the EURAXESS services can solve most of the problems regarding entry of foreign scientists into European research institutions and companies (*Nature* **488,** 557; 2012). Fast-track immigration is an important consideration for internationally mobile scientists and helps to attract the best global talent to Europe.

Ireland, like the United Kingdom and Denmark, was not obliged to introduce the scientific visa under the 2005 European Union (EU) directive, but it was recognized that the visa would help to attract non-EU researchers to the country. The scheme, which offers a free and fast service, has been successfully operating in Ireland since 2007 and is open to universities and companies.

Visas are issued rapidly and work permits are not required. Researchers' families can accompany them immediately and use public schooling. Family members have access to the job market and can stay on to look for a job after the contract ends.

More than 200 EURAXESS offices in 40 European countries, including Ireland, help foreign researchers to move across Europe by providing practical information. The EURAXESS portal (http://ec.europa.eu/euraxess) is a free advertising forum for EU research positions.

Conor O'Carroll *Irish Universities Association, Dublin, Ireland. conor.ocarroll@iua.ie*

Who discovered Universe expansion?

Controversy persists over who first found that the Universe is expanding. Last year, Mario Livio quashed suggestions that Georges Lemaître's 1927 theoretical prediction of expansion was deliberately suppressed (*Nature* **479,** 171–173; 2011). Since then, another contender has emerged.

The joint NASA and Infrared Processing and Analysis Center Extragalactic Database of Galaxy Distances, in Pasadena, California, which I co-lead, has tabulated and made public the historical distance estimates published by Edwin Hubble and his contemporaries to prove expansion (see I. Steer *J. R. Astron. Soc. Can.* **105,** 18–20; 2011). These reveal that measurements by a Swedish astronomer, Knut Lundmark, were much more advanced than formerly appreciated.

Lundmark was the first person to find observational evidence for expansion, in 1924 — three years before Lemaître and five years before Hubble. Lundmark's extragalactic distance estimates were far more accurate than Hubble's, consistent with an expansion rate (Hubble constant) that was within 1% of the best measurements today.

However, Lundmark's research was not adopted because it relied on one unproven method (galaxy diameters), cross-checked with one unproven distance to the Andromeda galaxy, which was derived from a type Ia supernova observed in 1885 and mistaken for a normal nova (W. Huggins and W. F. Denning *Nature* **32,** 465–466; 1885).

Hubble's research in 1929 yielded a value for the Hubble constant that was inaccurate by almost an order of magnitude. It was adopted because it was derived from multiple methods — including one still in use (brightest stars) — and was cross-checked with multiple galaxies with distances based on proven Cepheid star variables.

Lundmark established observational evidence that the Universe is expanding. Lemaître established theoretical evidence. Hubble established observational proof.

Ian Steer *NASA/IPAC Extragalactic Database of Galaxy Distances, Pasadena, California, USA. ian@colosseum.com*

Environment in Queensland at risk

In the state of Queensland, Australia, hard-won environmental protections are under threat.

In April this year, Queensland elected a new government that is pro-development and pro-mining. These activities have been burgeoning over the past few years, prompting the United Nations Educational, Scientific and Cultural Organization (UNESCO) to investigate whether the Great Barrier Reef should be included on the List of World Heritage in Danger (see http://whc.unesco.org/en/documents/117104).

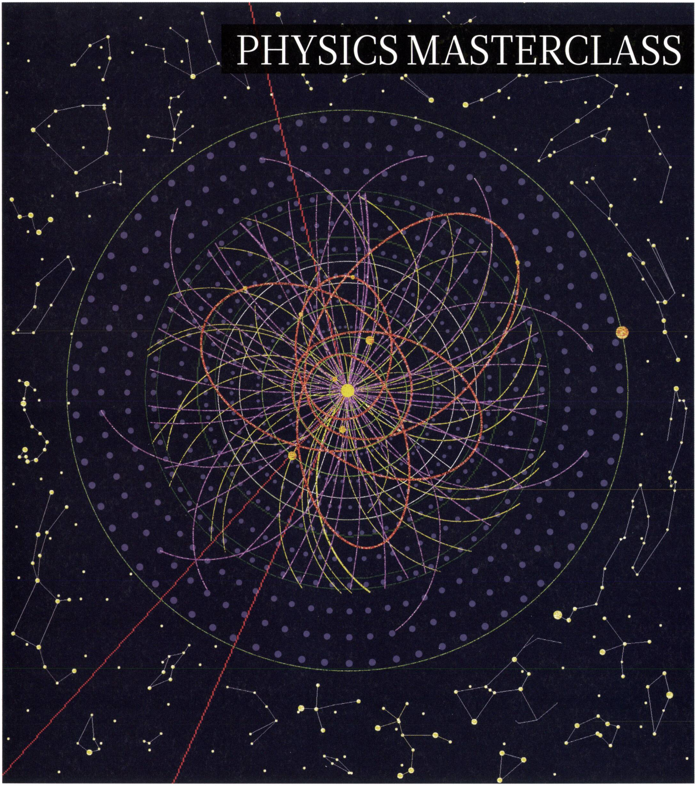
The government is also moving to dismantle the Wild Rivers Act, which was designed to protect and manage Queensland's last remaining pristine rivers and catchments in Cape York. These rivers have the richest diversity of freshwater fish in Australia, including the sawfish *Pristis microdon* and the speartooth shark *Glyphis glyphis*. These are considered by the International Union for Conservation of Nature to be critically endangered and endangered, respectively. Australia is one of the last countries in the world where populations of these fish are viable.

Ministers have nevertheless declared that a regional plan for Cape York will "unashamedly" fast-track development and "absolutely" secure a place for mining in these areas (see go.nature.com/e6hbzc).

These threats to Australia's aquatic environment are being compounded by government cuts to staff in environmental and fisheries management across public institutions.

Andrew Chin, Jimmy White *James Cook University, Townsville, Australia. andrew.chin@jcu.edu.au*
Nick Dulvy *Simon Fraser University, Burnaby, Canada.*

nature OUTLOOK

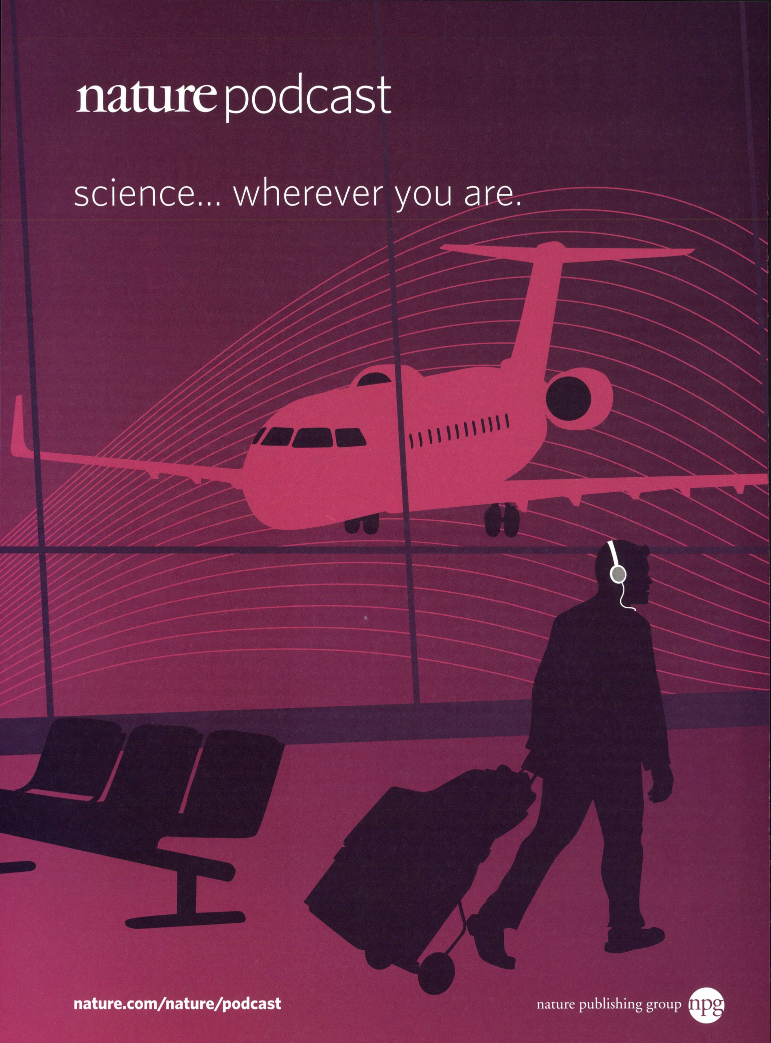


PHYSICS MASTERCLASS

Produced with support from:

MARS

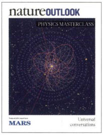
Universal conversations


nature podcast
science… wherever you are.
nature.com/nature/podcast
nature publishing group npg

natureOUTLOOK

PHYSICS MASTERCLASS

11 October 2012 / Vol 490 / Issue No. 7419



Cover art: Nik Spencer

Editorial
Herb Brody, Michelle Grayson, Tony Scully, Rebecca Dargie, Nick Haines

Art & Design
Wes Fernandes, Alisdair Macdonald, Gareth Richman

Production
Karl Smart, Kelly Hopkins, Leonora Dawson-Bowling

Sponsorship
Reya Silao, Yvette Smith

Marketing
Elena Woodstock, Hannah Phipps

Project Manager
Claudia Deasy

Art Director
Kelly Buckheit Krause

Magazine Editor
Tim Appenzeller

Editor-in-Chief
Philip Campbell

The annual Lindau Nobel Laureate Meeting always crackles with intellectual electricity. But the timing of this year's gathering — dedicated to physics — gave it an extra jolt. On 4 July, the 27 laureates and nearly 600 young researchers gathered on the beautiful German island of Lindau in Lake Constance watched a live announcement by CERN, the high-energy physics laboratory near Geneva, Switzerland, of the almost certain discovery of the Higgs boson. Many of the laureates at Lindau had been involved in laying the foundation for this discovery — David Gross was particularly enthusiastic, punching the air in triumph after the announcement — while other laureates and young researchers were huddled in groups, excitedly discussing its implications.

This *Outlook* celebrates the spirit of the Lindau meeting. The opening talk, by astronomer Brian Schmidt, described work on the dark energy that is driving the expansion of the Universe, for which he was awarded a share of the 2011 Nobel Prize in Physics. We look in more detail at dark energy — and what new tools might help probe its nature (page S2). We have Q&As with five of the laureates who attended the meeting, including theorist Martinus Veltman who, despite contributing to the mathematical framework that contains the Higgs boson, talks of his mixed feelings about its discovery (S10). And we present a conversation between astronomer John Mather, a 2006 prizewinner for his investigation of the cosmic microwave background radiation, which holds clues to the early Universe, and Minnie Mao, who recently took a position as a postdoctoral researcher at the National Radio Astronomy Observatory in Socorro, New Mexico. Mather and Mao discuss future strategies for observing our cosmic past (S5).

We are pleased to acknowledge the financial support of Mars, Incorporated in producing this *Outlook*. As always, *Nature* retains sole responsibility for all editorial content.

Michelle Grayson
Senior Editor, Supplements

CONTENTS

S2 COSMOLOGY
Out of the darkness
Something mysterious is driving the expansion of space

S5 ASTRONOMY
A conversation about observation
Two astronomers correspond

S8 Q&A
Douglas Osheroff
Finding fascination in nature

S10 Q&A
Martinus Veltman
Coming to terms with the Higgs

S11 Q&A
William Phillips
In search of ignorance

S13 Q&A
Roy Glauber
From Los Alamos to Stockholm

S14 Q&A
Mario Molina
Socially responsible science

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the Nature Outlook Editorial guidelines available at http://www.nature.com/advertising/resources/pdf/outlook_guidelines.pdf

CITING THE OUTLOOK
Cite as a supplement to *Nature*, for example, *Nature* **Vol XXX**, No. XXXX Suppl, Sxx–Sxx (2012).

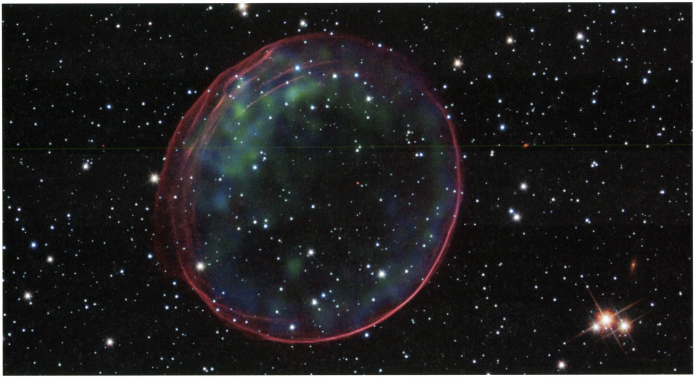
VISIT THE OUTLOOK ONLINE
The *Nature Outlook Physics Masterclass* supplement can be found at http://www.nature.com/nature/outlook/masterclass_2012.
It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES
For UK/Europe (excluding Japan): Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas – including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea:Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

CUSTOMER SERVICES
Feedback@nature.com

Copyright © 2012 Nature Publishing Group



Death of a star. Type 1a supernovae are the 'standard candles' of cosmology.

NASA, ESA, CXC, SAO, THE HUBBLE HERITAGE TEAM (STSCI/AURA), J. HUGHES (RUTGERS UNIVERSITY)

COSMOLOGY

Out of the darkness

The 62nd Lindau Nobel Laureate Meeting opened with a talk by Brian Schmidt, who shared the 2011 physics prize for the shocking revelation that the Universe is expanding at an accelerating rate. Fifteen years after Schmidt's initial discovery, the 'dark energy' invoked to explain this cosmic acceleration is still a mystery.

It is the most sobering fact in physics. Everything that has ever been observed, from elementary particles to clusters of galaxies hundreds of millions of light years across, makes up just 4% of the Universe. The remaining 96% is currently missing: dubbed dark matter (23% of the Universe) and dark energy (73%), their existence is inferred by the effect they have on ordinary matter, but their natures are unknown.

Astronomer Brian Schmidt, from the Australian National University's Research School of Astronomy and Astrophysics in Weston Creek, Australia, is partly to blame for revealing this unsettling situation. In 1998, his observations of exploding stars known as supernovae led to the conclusion that the Universe is imbued with something that is causing the expansion of space-time to accelerate. This wasn't what he expected to see: although it had been observed that the Universe was expanding, the prevailing idea was that this expansion was decelerating and that gravity would eventually stop the growth.

Schmidt recalls how he joked about possible outcomes before starting. "The worst thing that could happen to us is that we find that the Universe's expansion is accelerating because people would just say that we were crazy." Yet, in 1998, this is exactly what they did see. What's more, he adds, the observations suggested this dark energy could be a form of antigravity, "having physical properties that essentially makes gravity repulsive instead of attractive — it's a pretty big deal."

The discovery was made jointly by two groups who shared the 2011 Nobel Prize in Physics: the Supernova Cosmology Project, led by astrophysicist Saul Perlmutter of Lawrence Berkeley National Laboratory and the University of California, Berkeley (awarded half of the prize), and the High-Z Supernova Search Team led by Schmidt and Adam Riess of Johns Hopkins University in Baltimore, Maryland. Schmidt and Riess's half was split between them.

Together with ordinary matter and radiation, dark energy and dark matter are the basic ingredients of the standard model of cosmology that accounts for 13.7 billion years of cosmic evolution. Dark matter might one day be revealed by experiments — one candidate, 'supersymmetric' particles, are being searched for by physicists at CERN's Large Hadron Collider, for instance. Dark energy, on the other hand, presents a new kind of bafflement. "You have no intuition or guidance because the physics is so different to anything we have experienced," says cosmologist Robert Caldwell of Dartmouth College in Hanover, New Hampshire.

RELATIVE PREDICTIONS

The framework for the standard model of cosmology is provided by Einstein's 1916 general theory of relativity, which regards gravity as a consequence of matter causing space-time to curve. The general theory predicted a dynamic Universe that could either expand or contract. Since scientists at the time believed the Universe to be static, Einstein inserted an extra term into his equation to enable his theory to fit with convention: a 'cosmological constant' called lambda. Lambda was generally abandoned a decade later, after astronomer Edwin Hubble and others showed that the Universe is in fact expanding and not constant. But perhaps Einstein was on the right track, even if he did not know it: general relativity is consistent with recent observations if the cosmological constant is reinstated with, as Schmidt puts it, a "ridiculously small value" of roughly 7×10^{-27} kg m^{-3}. (By comparison, the density of the air around us is about 1.2 kg m^{-3} — a billion billion billion times greater.)

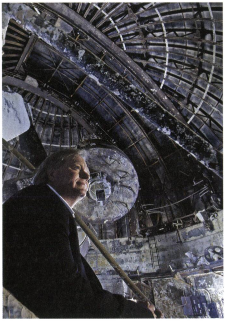
This tiny number has big consequences for cosmic evolution because, unlike matter or radiation (whose density decreases as the volume of space increases), lambda has the same value regardless of the size of the Universe. Consequently, if dark energy is equivalent to lambda, its contribution to the total density of 'stuff' in the Universe will be greater in a bigger universe, and vice versa. This is important because the relevant amount of stuff determines whether the expansion of the Universe will eventually grind to a halt or continue for eternity.

Data suggest that the Universe passed a tipping point about 6–7 billion years ago (see 'A universal lifecycle'). Prior to this point the Universe was dominated by matter — visible and dark — which was slowing the expansion. Then dark energy took over, driving an ever faster — perhaps exponential — expansion. So instead of the Universe gradually winding down under gravity, dark energy propels it to a far different fate, explains Caldwell. He says we could be headed towards a 'Big Chill' in which our local group of galaxies becomes an island in an eternally cold and lonely Universe, or possibly even a 'Big Rip' — a catastrophic demise that would "pull apart the Milky Way, the Earth, and ultimately atoms and nuclei".

The cosmological constant is so small, yet not quite zero, that to explain it some cosmologists are entertaining the idea that our Universe may be part of an infinite number of universes, each encoded by different fundamental laws and causally decoupled from one another. In this picture the 'unnaturally' small value of our cosmological constant is statistically plausible given the large number of other possible universes out there — and we just happen to be in this one. If its value were much different, galaxies, stars and people would not exist. Such anthropic reasoning is not to everyone's taste. "It could turn out to be the best way to explain dark energy," says Perlmutter, "but it's a lot to take on board unless it makes some kind of prediction that we can test."

IMAGINEERING THE UNIVERSE

Cosmologists are in need of a new theoretical insight — a "big leap", says Schmidt. Since 1998, more than 7,500 papers with "dark energy" in the abstract have been published — most of which tackle its theoretical roots. Within these theories, a cosmological constant is not the only solution proposed: one alternative invokes a new type of field called quintessence (from the classical name for the fifth element beyond earth, air, fire and water). Rather than remain constant, quintessence allows dark energy to vary over time and space. It is described as a scalar quantum field — an entity associated with an elementary particle that has no spin — defined by a simple value at all points in the Universe.



The Mount Stromlo Observatory, where Brian Schmidt (pictured) made his observations, was destroyed by bushfire in 2003.

ANDREW CAMPBELL/FAIRFAX SYNDICATION

Scalar fields are not new ideas: they are also suspected to have caused an extremely short-lived period of expansion of the Universe called inflation, when space expanded by a factor of 10^{80} from approximately 10^{-35} to 10^{-32} seconds after the Big Bang. Inflation would have amplified the tiny events that are naturally part of space-time (quantum fluctuations) to macroscopic scales, generating regions of differing densities, which allowed gravity to kick in and cosmic structure to form. Until recently, scalar fields were the stuff of imagination. But CERN's announcement in July 2012 of a new particle suspected to be the Higgs boson (a scalar particle) suggests that space is filled with at least one scalar field — the Higgs field that gives elementary particles their mass.

> ***"I don't think that we've learned anything significant about dark energy since 1998."***

"Cosmologists have been working with fundamental scalars for 30 years without knowing if they exist," says astrophysicist Pedro Ferreira of the University of Oxford in the United Kingdom. "So if this new particle is a scalar then it's a big deal." Ferreira adds, however, that the energy scales of the Higgs field, dark energy and inflation differ by so many orders of magnitude that it is difficult to build convincing theories linking the phenomena.

However, scalar fields are only one option. "There are hundreds if not thousands of possibilities," says Ferreira. Dark energy might be a vector field — one that has both a magnitude and a direction at each point in space. Other, wilder theories postulate higher dimensions of space. But Ferreira laments the overall lack of progress. "I don't think that we've learned anything significant about dark energy since 1998."

Some cosmologists don't even think dark energy is real: rather, it is a consequence of two assumptions that underlie the standard cosmological model. The first assumption is that Einstein's general relativity holds everywhere in the Universe; the second is that the Universe is homogenous and isotropic. So far, both assumptions have been supported by observations, yet at distant regions of space they might break down.

One doubter is theorist Subir Sarkar at the University of Oxford: "I don't dispute the measurements, but I think dark energy is an artefact of interpreting data in the simplest possible cosmological model — formulated when there were essentially no data available to test homogeneity." Sarkar argues that the data show only that distant supernovae are fainter than expected for objects at that redshift (the change in colour of light from supernovae as space-time stretches). It is another assumption, however, to say that that region of the Universe is undergoing the same rate of expansion as here. "The expansion rate would vary in space — not just in time — in an inhomogenous Universe."

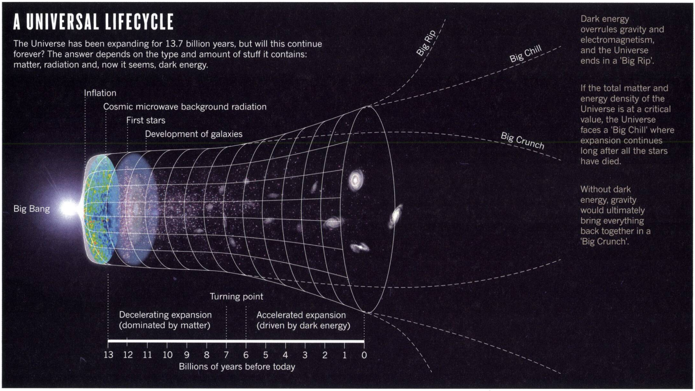
It could be that general relativity needs some tweaking when it comes to describing the largest scales. Physicists have grappled for some time with modified-gravity theories as alternatives to invoking dark matter, but the accelerating expansion of the Universe adds another experimental observation for theories to accommodate. "We are in a mess with dark energy," says Sarkar. "Astronomers like the simplest solutions of general relativity but there are other, more complicated solutions that we have to consider."

Dark energy may not be an attractive proposition, admits Schmidt, but with it the standard model of cosmology correctly predicts the geometry of the Universe, temperature fluctuations in the cosmic microwave background radiation (CMBR) and the abundance of light elements forged in the Big Bang. "The only problem," Schmidt notes, "is that it requires us to invent 96% of the Universe."

FINDING DARK ENERGY

Pending new theoretical insights, researchers are badly in need of more data. Upcoming experiments will refine measurements of the acceleration. Since 1998, the precision of the supernovae data has been improved tenfold, says Schmidt. But measurements are now at the level in which systematic errors, such as the ability of astrophysicists to measure how bright things are, dominate statistical errors caused by the finite number of objects observed. Knowing the brightness of supernovae is vital because these objects are considered 'standard candles' with which to infer cosmic distances.

The "big one" for dark energy, says Schmidt,



A UNIVERSAL LIFECYCLE

The Universe has been expanding for 13.7 billion years, but will this continue forever? The answer depends on the type and amount of stuff it contains: matter, radiation and, now it seems, dark energy.

Inflation
Cosmic microwave background radiation
First stars
Development of galaxies
Big Bang
Turning point
Decelerating expansion (dominated by matter)
Accelerated expansion (driven by dark energy)
13 12 11 10 9 8 7 6 5 4 3 2 1 0
Billions of years before today
Big Rip
Big Chill
Big Crunch

Dark energy overrules gravity and electromagnetism, and the Universe ends in a 'Big Rip'.

If the total matter and energy density of the Universe is at a critical value, the Universe faces a 'Big Chill' where expansion continues long after all the stars have died.

Without dark energy, gravity would ultimately bring everything back together in a 'Big Crunch'.

is the European Space Agency's (ESA's) Euclid mission, which in June 2012 got the go-ahead for construction, to launch in 2019. This satellite will use visible-wavelength and near-infrared cameras to map the three-dimensional distribution of more than one billion galaxies stretching over distances of tens of billions of light years. Euclid will measure so-called baryon acoustic oscillations — periodic clumpiness in the distribution of matter caused by sound waves that propagated in the hot plasma of the young Universe. In this way, the mission aims to track how the expansion of the Universe has changed over time.

Euclid will also exploit a technique called weak gravitational lensing, the bending of light from faraway galaxies as it passes through the gravitational field of nearer matter. Like acoustic oscillations, gravitational lensing is sensitive to the structure of matter in the Universe over time. Weak gravitational lensing tells cosmologists how quickly structure formed and so allows detailed tests of general relativity on the largest scales. "Modified gravity used to be a fringe thing, but it has now become an integral part of missions such as Euclid or SKA [the Square Kilometre Array radio telescope]," says Ferreira.

NASA has proposed a similar probe to Euclid called the Wide-Field Infrared Survey Telescope (WFIRST) , designed to give researchers a better handle on the brightness and spectral evolution of supernovae. Should the project get the go-ahead, WFIRST would measure the properties of more than 1,000 supernovae at various distances. Meanwhile, NASA's James Webb Space Telescope (JWST), the successor to the Hubble Space Telescope, is due to launch in 2018 (see 'A conversation about observation', page S5). In 2004, it was Hubble's high-redshift (or deep-field) observations of supernovae that enabled Riess and colleagues to pin down the time at which the cosmic acceleration kicked in; the JWST will take this further. Cosmologists are also looking forward to tighter constraints on the geometry of the Universe from recent measurements of the CMBR taken by ESA's Planck satellite.

DARK FUTURE

Shortly after the 1998 result, recalls Perlmutter, his group was discussing what it would take to obtain sufficient data for theorists to gain an understanding of dark energy. "We knew we had a decade of development work to do, and now that we have done that we are ready to build these projects," he says. "If we can spend the next decade measuring things then I am optimistic that we will make progress in understanding dark energy. This is a fairly new field, so there is a reasonable chance of a breakthrough."

First, Perlmutter and others will take Earth-based measurements. These include the Dark Energy Survey, a specially adapted camera fitted to the 4-metre Blanco telescope in Chile, which from this year will use supernovae, baryon acoustic oscillations, weak lensing and measurements of large-scale structure from hundreds of millions of distant galaxies. Weak lensing is also being exploited by an instrument fitted to the 8.2-metre Subaru telescope in Hawaii. Towards the end of the decade, an experiment funded by the US Department of Energy will study baryon acoustic oscillations using one of the National Optical Astronomical Observatory 4-metre telescopes — a project called BigBOSS is a leading contender.

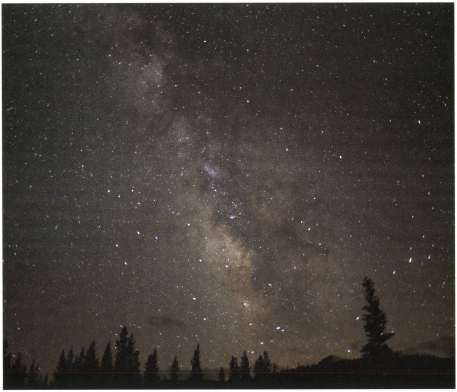
Meanwhile, projects at telescopes in across the world — for example the Skymapper telescope in New South Wales — will give astronomers a better idea of the brightness calibration of supernovae in the nearby (low-redshift) Universe. Further in the future is a high-resolution optical spectrograph called Codex, which will aim to provide an absolute measure of the cosmic expansion rate (enough to convince even the hardest of acceleration sceptics). Codex will be installed on the recently approved European Extremely Large Telescope in Chile and will measure the change in the Hubble expansion rate — the 'redshift drift' — over a ten-year period.

The data will drive new theories, researchers hope. "In terms of understanding the rhyme or reason of how it works, there have been many papers on dark energy but no breakthrough ideas," says Craig Hogan of the Fermi National Accelerator Laboratory in Batavia, Illinois. "As far as experimental prospects go, there are experiments that can test extensions of standard physics, if not dark energy directly, that may point towards some emergent theories of cosmic acceleration."

They need to seize the moment: if the expansion of the Universe is accelerating, then in a few tens of billions of years everything will have disappeared over the cosmic horizon and be too far away for its light to ever reach us. "We are losing information by the second," says Schmidt. "The number of atoms that are in our horizon is dropping rapidly and has been for about 6 billion years, so it's an interesting time that we live in." ■

Matthew Chalmers *is a freelance science writer based in Bristol, UK.*

ISTOCKPHOTO



ASTRONOMY

A conversation about observation

In the spirit of the Lindau Meeting, we present a dialogue between a Nobel laureate and a young researcher. This interchange started online, where it continues to unfold. Here is a digest of this conversation, which has developed across time and space.

MEET THE PARTICIPANTS

Two astronomers correspond

John Mather shared the 2006 Nobel Prize in Physics for his work on the cosmic microwave background radiation. Mather is currently at NASA, focusing on building Hubble's successor, the James Webb Space Telescope (JWST), which he spoke about at the 2012 Lindau Meeting. The JWST is due to launch in 2018 and will orbit the Sun about 1.5 million kilometres further out than the Earth's orbit (in a Lagrange point) to observe the very first galaxies. **Minnie Mao** recently took a position as a postdoctoral researcher at the National Radio Astronomy Observatory (NRAO) in Socorro, New Mexico — leaving her native Australia. Minnie uses radio and optical telescopes to study the evolution of galaxies to determine how they have formed and evolved since the Big Bang 13.7 billion years ago.

"I'm trying to figure out what my role is in the world of astronomy."

Dear Professor Mather,

It was really lovely to meet you at the Lindau Meeting. I had a wonderful, thought-provoking week that left me feeling buzzed up and raring to do science!

My name is Minnie and I've just started a postdoc at the NRAO. I submitted my PhD on the cosmic evolution of radio sources in April and will (hopefully!) graduate from the University of Tasmania in December this year. New Mexico is somewhat different from Tasmania, but I'm really enjoying it here — especially being based at the array operations centre for the Karl G. Jansky Very Large Array (VLA).

This brings me to why I do astronomy: I love observing. I love the process of deciding to 'look' at something, setting up the telescope, waiting in breathless anticipation for the photons to hit the detector, or watching the cross-correlations dance across a computer screen. I even love reducing the data, knowing that shortly I will be seeing something no one has ever seen and using it to answer my science questions.

What do you think will be the most important result to come from the JWST? Will they be driven by predicted theories or by observations that yield completely unexpected results, which we couldn't possibly have known had we not observed something new, as was the case with the Hubble Deep Field?

I love my new job. Learning all about the VLA is so much fun, but now that I'm starting to get used to the fact that I'm no longer a student, I'm trying to figure out what my role is in the world of astronomy.

Clear skies,
Minnie Yuan Mao

"After all these years, we still have a hard time making a theoretical supernova explode."

Dear Minnie,

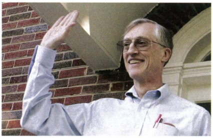
Thanks for your letter! I like to be called John; I'm still not used to formal titles.

My own passion for physics and astronomy has a lot to do with a desire to build things. When I was a teenager I learned a lot from the *Scientific American* monthly feature called *The Amateur Scientist*, and I borrowed the 3-volume book series *Amateur Telescope Making* from the library over and over again. I did build a small telescope, but

I didn't have the patience to grind a mirror — I bought that. I also loved thinking about electronics, and I wanted to be a ham radio operator, so I assembled a Heathkit AR-3 radio receiver (with vacuum tubes) and read *The ARRL Handbook* from cover to cover, over and over.

There's a wonderful book by Martin Harwit called *Cosmic Discovery* that lists the greatest discoveries in astronomy, and shows what made them possible. In almost every case, the discovery came from new equipment, and in a pretty large fraction of cases the equipment was designed for some other purpose. So I see a path to great discovery: imagine a measurement that could answer a question, and build something to make the measurement. Then, either you answer the question, or you find something new. But in either case, it will be something nobody else could do. Have you given any thought to good strategies for making your own discoveries?

That leads to your other question: what will be the most important discoveries with the JWST? Of course, I don't know specifically, but it seems that surprise is part of importance — if everyone knows something already, it is not a discovery. No theorists or observers predicted the correct story for galaxy formation before the Hubble Space Telescope was launched, neither group predicted cosmic dark energy or dark matter (though Fritz Zwicky saw evidence for dark matter in the 1930s), and neither predicted the nature and abundance of planetary systems beyond our own Solar System. My friends doing numerical simulations of stars and galaxies in supercomputers agree that they are not going to predict what we will find with any certainty — the complexity is just too great. And after all these years, we still have a hard time making a theoretical supernova explode.



NASA's John Mather says discoveries come from new equipment.

LARRY DOWNING/REUTERS/CORBIS

So, I think the JWST will lead to some great surprise that nobody is even thinking about today. Perhaps, you will be the one to find that surprise.

With best wishes for your new role as a Discoverer!
John Mather

"Learning is not just about accumulating information; it's also about the process of getting that information."

Dear John,

I like the 'discoverer' tag: I think that is what science is about — the discovery of new things, new phenomena, new concepts, new ideas.

VLA VISITOR

Minnie Mao believes in pushing a telescope's capabilities.

M. MAO

You asked if I'd given any thought to good strategies for making my own discoveries. I've thought about this a lot. When I was applying for postdocs at the end of last year I looked for jobs that had a telescope support element. I wanted a job that would let me get to know the telescope. I figure that understanding a telescope's capabilities and limitations will enable me to ask science questions that the telescope is able to answer.

I did my PhD in Australia where I spent a lot of time at the Australia Telescope Compact Array (ATCA). I learned so much by being on-site, things like how much microwave ovens affect observations at 13 cm (a lot!). Experiences like these enabled me to understand my data intimately and hence distinguish real science from image artefacts and other distractions. As such, I felt like I was able to tailor my projects to make the most of the ATCA's capabilities.

Conversely, the VLA has dedicated telescope operators who observe for you. I wish I could do more of the observing myself though. I think learning is not just about accumulating information; it's also about the process of getting that information. I could go away and read all the how-to manuals about the telescope and its data products, but until I actually see the telescope work, and reduce the data myself, I will not truly understand it.

I feel like I'm at the tail-end of the generation of astronomers that gets to actually play with telescopes. Hands-on observing is fast becoming obsolete. What will happen to the next generation of astronomers who write proposals for telescope time? We're always going to have people churning out good, solid science. But is anyone going to have the expertise to come up with something truly innovative?

Unless you are directly involved in the building and commissioning of a new telescope, do you think you can really understand it enough to design creative experiments that will produce ground-breaking science?

Clear skies,
Minnie

"The challenge is to use our imagination to the fullest."

Dear Minnie,

I like your strategy of 'playing with telescopes' to understand their capabilities and limitations. I think it's important because so many great discoveries happen at the very edge of capabilities. The amazing story of the early galaxies is only barely discernible from tiny coloured dots on the Hubble deep-field images, and it has taken tremendous persistence on the part of the observers and instrument builders to know how to: (a) get scientific discoveries from little coloured dots; and (b) get better equipment to see them more clearly. Similarly, what we know of exoplanets has come mostly from people who knew how to push the capabilities of existing equipment (first), and then pushed to get better equipment (after the richness of the discoveries became apparent).

You seem a bit concerned that you won't be able to learn enough about telescopes without touching them. I have that feeling too, but I think we are more and more able to experience things we can't see or touch. After all, astronomy is all about things we can only see from immense distances! I'm writing this on a computer running on hardware I've never seen, based on quantum mechanics that hardly anyone understands. At NASA, I'm working on a giant telescope that for years existed only in the imagination, along with a few documents and viewgraphs, and now two of the four instruments are delivered and all of the mirrors are finished.

So, the challenge is to use our imagination to the fullest! Hands-on observing is a way to build that imagination, but total immersion in using something you can't touch still works. So to answer your question, yes, I truly expect there will be many people with the expertise to come up with something truly innovative. We are building the JWST so that a generation of people who were not even born when it was started will be able to use it for breathtaking discoveries.

Fear not, just dive in!
John

“What I do object to is the complete removal of the astronomer from the observing and data-reduction process.”

Dear John,

I’m really pleased to hear that you’re not worried about the reduction in hands-on observing opportunities. While hands-on observing is something I enjoy passionately, I guess I can understand that it’s not necessarily vital to being a brilliant astronomer.

Having said that, I disagree that imagination can replace the wealth of knowledge garnered from sitting in front of a terminal and actually being in control of a telescope. You say “total immersion in using something you can’t touch still works”, but is that really true? To use your computer analogy, you can read hundreds of books about computers, but you learn most about how to use your computer by sitting at it and actually playing with it.

The Australia Telescope Compact Array.

ANGEL R. LOPEZ-SANCHEZ/AUSTRALIAN ASTRONOMICAL OBSERVATORY/MACQUARIE UNIVERSITY

In my last letter, I focused on being physically at a telescope to use it. Obviously this is not feasible for telescopes such as Hubble and the JWST. But even in these circumstances, what I do object to is the complete removal of the astronomer from the observing and data reduction process. Radio telescopes are moving towards a model where the astronomer puts in their observing request and is then provided with completely reduced data. I understand that this is somewhat necessary due to the exorbitantly large sizes of the data sets, but I still find it troubling.

Twenty years ago if you wanted to use a telescope, you went to the telescope, you observed, and you were responsible for reducing your data and analysing it. Nowadays, hands-on observing is becoming less prevalent. Science isn’t suffering yet because the astronomers still retain their observing and data-reduction experience. But what happens 20 years from now, with a new generation of astronomers? I’m not convinced that truly innovative ways of exploiting a telescope will be possible when the telescope users are oblivious to the observation and data reduction processes.

What are you planning on doing with the JWST? Will approved observing programs deliver completely reduced data? Will the astronomer even know when their proposed objects are being observed?

Clear skies,
Minnie

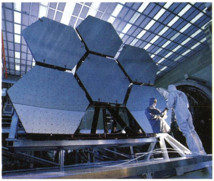
“There is never the final answer to anything, only the best answer we think we have at the time.”

Dear Minnie,

I agree completely that hands-on observing is the way to go when possible. There’s nothing like actually touching things and studying them in person to get acquainted. One wouldn’t (I hope) marry somebody over the internet without meeting him/her! But as time goes on we all get used to ignoring the inner workings of things, as long as they are reliable. Hardly anyone even knows how a car works these days, or even suspects that it has perhaps a hundred microprocessors running little dedicated tasks. So we get to concentrate on top-level issues, like where to go and how fast. I think modern astronomy has gotten to that point too — we have software packages with millions of lines of code, and we trust that somebody has made them work right. And we have observatories with complexities that are completely hidden from the users most of the time. Of course, the sceptical and cautious user (a real scientist) will not believe everything in the users’ manual for either a telescope or a software package, and will find a way to test every important feature, otherwise we will have a lot of non-science. I think the real breakthrough science will still come from people who push the capabilities of these tools beyond their design limits and test what happens. That will take familiarity with design, and imagination to picture what cannot be seen.

About JWST, there’s no such thing as completely reduced data from any observatory. We do expect that the observers will receive data already converted into units of astronomical brightness, according to the best calibration information and algorithms we have. But if an object is very faint, it will have to be observed over a long period of time with multiple exposures that have to be combined; the process of doing the combination is not so simple, and the cautious observer may want to do it him/herself. But that’s still a very long jump from data gathering to interpretation of the objects way out there — that remains the observer’s remit. And there is never the final answer to anything, only the best answer we think we have at the time. And yes, the astronomers will know when their proposed objects will be observed. (Small caveat: we are setting up the telescope to have a little autonomy, so if an observation cannot be completed in sequence, the telescope will automatically go to the next one on its list.)

➲ NATURE.COM
See Mather debating The Golden Age of Astronomy
go.nature.com/Luk8pf



The recently finished mirrors of the James Webb Space Telescope.

NASA/MSFC/DAVID HIGGINBOTHAM/EMMETT GIVEN

I find it very exciting to know of the discoveries made by people working far from my own specialties. And once in a while a conversation occurs that starts a new field of research, just from asking a question about something that is obvious to one person but not to another. We have dark matter in astronomy that is a total mystery. But there is a form of dark matter in genetic material too, and just lately it was announced that it is extremely important: the dark matter between the genes is actually millions of digital switches. I don’t know when a breakthrough like that for dark matter and dark energy will occur for astronomers. But why not? And in another area of astronomy, suddenly we know that there are as many exoplanets as there are stars, that small ones are very abundant, and that many have been expelled into space and have no star any more.

So I think one of the most interesting ways to phrase a question in science is: “If X happens, then what?” What do you think are interesting questions?

All the best,
John

To see this conversation continue online, visit lindau.nature.com/lindau/online dialogue

A handful of Nobels

Five laureates from the 2012 Lindau Nobel Laureate Meeting give us their thoughts on everything from the Higgs boson to climate change.

CHINAFOTOPRESS/GETTY IMAGES

Q&A DOUGLAS OSHEROFF
Finding fascination in nature

Douglas Osheroff is a Stanford University physicist focused on ultralow-temperature physics of quantum fluids and solids. Co-recipient of the 1996 Nobel Prize in Physics for the discovery of superfluidity in helium-3, Osheroff spent much of his childhood performing science experiments — often with an element of danger — that helped hone his technical skills.

You were interested in natural sciences from an early age. What triggered this?
I was six when I got an electric train for Christmas. My parents took a nap in the middle of the afternoon; by the time their nap was over I had completely disassembled the locomotive to get the electric motor out. I don't know what made me do that!

From age five to when I graduated from high school, I was fascinated by electricity and magnetism. A lot of the stuff I did was dangerous, although my parents didn't reprimand me or try to force me into safer activities. I think that my father was actually intrigued by my fascination with things electrical.

Later, my high-school chemistry teacher spelled out the way he thought about research. He said that when you do research you ask questions of nature. That idea has always fascinated me, and in fact describes the way I approach research. Every time you do an experiment you are asking a question or a series of questions of nature, and nature then has to answer them — but nature's answers are complicated, so you have to ask more questions.

Did you have any mentors, and how did they shape your academic development?
My father had a lot to do with developing my interest in science. Other than that, Richard Feynman had an impact on me. In my freshman class, everyone at Caltech (California Institute of Technology) had to take two years of Feynman physics. For people like me who were absorbed by nature, that was phenomenal, because he took his fascination and combined it with mathematical understanding in a way that made physics a more interesting and more powerful science.

How did you become interested in low-temperature physics and helium-3?
My interest with low temperatures came about in primary (elementary) school when I started freezing mercury. Our class would go on picnics and my mother would always provide frozen desserts that were packed in dry ice. At the time, I also had a lot of mercury that I had taken out from used batteries collected from drugstores around town. When I put the mercury on this dry ice it froze, it became a solid. I was fascinated.

But my interest really peaked when I was doing graduate work at Cornell University at the end of the 1960s. There were two talks during the first semester that I attended on new refrigeration devices that had been developed very recently. I recognized that these new refrigeration technologies would allow me to look at nature in a different way, and see things that no one had ever seen before.

One was the helium-3–helium-4 dilution refrigerator — in fact I built one in my first year of grad school. The other was Pomeranchuk cooling, which is a very strange thing. If you take liquid helium-3 at sufficiently low temperatures, just above absolute zero, and you compress it under a pressure of 34 atmospheres, you start converting the liquid to solid. For every per cent of the liquid you convert to solid, the system cools by 0.001 K. Now that does not sound like very much, but if you start at a temperature of 0.015 K, you only have to convert 15% of your sample to solid to reach absolute zero. Of course, you don't reach absolute zero, but you can get pretty close. When I heard the talk on Pomeranchuk

> **"Every time you do an experiment you are asking questions of nature, and nature has to answer them."**

EMILIO SEGRE VISUAL ARCHIVES/AMERICAN INSTITUTE OF PHYSICS/SCIENCE PHOTO LIBRARY

cooling, I said: "I have to have one of these things." So I started building a Pomeranchuk refrigerator.

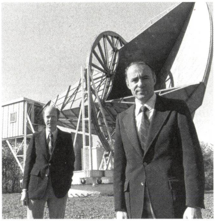
I didn't know what was likely to happen at temperatures this close to absolute zero. When I was testing my Pomeranchuk cooling device, all of a sudden I saw a sharp decrease in the rate of cooling. I was unhappy to see this and, at first, I tried to resume the original rate of cooling. But ultimately I found that this decrease was a signature of some highly reproducible phase transition in the mixture of liquid and solid helium-3. It took me a while to understand, but the signature was due to a jump in the heat capacity as the liquid became a superfluid. To this day (even right now as I tell you the story), the hair stands up on the back of my neck because I remember the excitement when I realized what I had discovered.

What are the key elements leading to the successful conception and execution of Nobel-prizewinning research?
If you're an experimentalist, you have to know how to design and build things, and understand what's going to work and what won't. At the same time you need to have a deep curiosity about the behaviour of matter and its limits — in my case at the limit of very low temperatures. And then you need to have a good understanding of what you expect to see, because what is probably most important is seeing something that you didn't expect — and you cannot know what that is unless you know what you expect in the first place. Of course, there is another side to this — when you clearly don't know what to expect, but are able to realize the importance of your unexpected findings, as was the case with Arno Penzias and Robert Wilson in their discovery of cosmic background radiation.

ROBERT HOISCHEN

You (and Penzias and Wilson) did groundbreaking research while at Bell Labs during its golden era. What were the key ingredients for Bell Labs' research successes, and do you think that similar research environments exist or could be created today?
Things have changed a lot. Bell Labs is a very different place than when I was there. Then it was a very big and complicated organization. I was in Area 10, and we were allowed to research whatever we wanted. The basic research part of Bell Labs was supported by funds that (parent company) AT&T got from the regional operating companies. Once the Bell System was broken up in the 1980s, that funding disappeared. I think that was the beginning of the end of fundamental scientific research at Bell Labs, although applied research continued.

It's almost impossible to create a similar environment today. Almost all research laboratories are supported by companies that have to show profit. If you look back to the 1970s, there were a lot of industrial research laboratories: General Motors had a lab, General Electric had one. At some point these companies began to realize that these laboratories weren't giving them any competitive advantage. The leaders of AT&T realized that basic research at Bell Labs is fun, and interesting things come out of it, but questioned whether the Bell System profited from it.



Bell Labs physicists Robert Wilson (left) and Arno Penzias, in front of the antenna that first detected the cosmic background radiation.

Pressure to publish is felt keenly by researchers, especially at the start of their career. What publishing strategies would you recommend to aspiring scientists?
It's tricky; the 'publish or perish' attitude might encourage people to publish things that aren't very exciting, or to put a little spin on it, so that's something to be careful about. Some people don't want to publish something that they don't completely understand, but that can be very limiting. For one thing, you won't publish much. Plus, there is a tremendous value to getting your results out to the scientific community, both so that your work can be evaluated but also so that someone else can look at your results and potentially spot something important that you've missed. It is important that results be shared.

If you could give your pre-university self one piece of advice, what would it be?
The same advice that I give my undergrad students at Stanford. You need to understand what your motivation is, who you are, what you enjoy doing, and what you're good at. If you optimize all those things, you will be a good scientist.

Have you forgiven the Nobel committee for waking you up at 2:30 in the morning?
[Laughter] It's really interesting. There have been times when I thought maybe my life would be richer and fuller if I had not got the Nobel prize. But you end up doing different things. One has to be careful though. There are a lot of organizations that particularly love having Nobel laureates on their advisory boards, not because the laureates are going to give good advice, but because it looks nice to have them on your advisory board. I am now at a point where I'm pretty good at saying no to this sort of thing.

There are exciting things that people invite you to do, mostly because you have the Nobel. These opportunities have distracted me from what I used to do; ironically the Nobel prize resulted in my doing less science!

Some people are able to put all of this stuff behind them and continue researching. In particular, in high-energy physics, where there's a lot of money involved, researchers can use the Nobel prize as a lever to get resources to do experiments that they couldn't do otherwise.

But it's a fascinating life. I visit all sorts of different places. I give talks that have important messages in them, and I hope that some people recognize the importance of the strategies I describe. I will continue doing what I'm doing until I'm too old to do it. ■

Interview by **Dan Csontos,** *a science writer based in Lund, Sweden.*

Question posed by Ulrika Forsberg, a PhD candidate at Lund University in Sweden, who is studying superheavy elements.

It's reassuring to know that good practices in the basics of science might actually help you do Nobel prizewinning research! But, although obvious, these things are not necessarily easy. For example, to understand what's going to work and what won't requires a thorough knowledge of the field of research and the experimental technique. This makes Osheroff's answer not so much a checklist on how to succeed, but rather hints that persistence as well as curiosity are what brings you to your goal.

SCIENTIFICAMERICAN.COM
You can visit Ulrika's 30 Under 30 page: go.nature.com/zaD4ry

Q&A MARTINUS VELTMAN

Coming to terms with the Higgs

This year's Lindau meeting coincided with the biggest particle-physics discovery in a generation. Theoretical particle physicist Martinus Veltman, emeritus professor at the University of Michigan in Ann Arbor, shared the 1999 Nobel Prize in Physics for his work on the 'standard model' of particle physics — the theory that predicted the Higgs boson. Yet he has spent the past 30 years doubting whether the Higgs exists.

What, basically, is the Higgs boson?
There is really no classical picture. The quantum field theory of weak interactions (which is comparable to, say, electromagnetic interactions) shows mathematical difficulties that can be traced to the non-zero mass of the intermediate particles, the W and Z bosons. A solution to this problem was to suppose the existence of a field extending through the Universe. Then the intermediate particles could interact with this 'Higgs field' and acquire energy, that is, mass. In quantum field theory, whenever there is a field there is an associated particle, for example the electromagnetic field has the photon. While the photon is massless and has spin 1, the Higgs boson is massive and has no spin.

What does the discovery of the new particle mean to you?
I reacted rather coolly at first, but now that I've had some time to reflect I've found myself getting wrapped up in what it means. The prediction of this particle was made 50 years ago, and now we have found it. It's really a very amazing consequence of the theory. The Higgs is a particle that should be there for rather abstract theoretical reasons; through the years I have investigated the consequences of it not being there. The discovery means I will abandon this effort, at least for a while.

Why did you ever doubt its existence?
First, this animal was supposed to be everywhere, giving mass to the other particles, yet it had escaped detection for 30 or 40 years. Second, since the energy of the Higgs is distributed all over the Universe, it should contribute to the curvature of space; if you do the calculation, the Universe would have to curve to the size of a football. That's one of the biggest problems in particle physics.

Are you now a convert?
It is not a matter of belief and emotion. I just look at a situation and I wonder: what is the evidence? The hunt for the Higgs started in the early 1970s, and every year there would be another joker predicting that it is just around the corner. I concluded that the chances of it not being there were higher.

Fortunately I got a call before the CERN announcement to tell me the result, just to prevent me saying something stupid at Lindau like "the Higgs doesn't exist"!

What role did your work play in predicting the Higgs boson?
My research concerned the mathematical formalism that you need to understand the 'renormalizability' of a certain class of quantum field theories including the electroweak standard model. Before that, you couldn't make a single prediction because you didn't know how to calculate things. It's similar to what Richard Feynman and others who got the Nobel prize did with quantum electrodynamics in 1948.

Did you ever work with Feynman?
We never published together, but once in 1983 we went for a hamburger and shared our dislike of supersymmetry.

Won't supersymmetry explain dark matter?
Of course it won't. People have been looking for this stuff since the 1980s and are just talking ballyhoo. Isn't it more likely that we don't understand gravity all that well? Astrophysicists believe in Einstein's theory of gravity with a fervour that is unbelievable. Do you know how much of Einstein's theory has been tested at the distances of galaxies where we 'see' dark matter? None of it!

*** Feynman said that if you couldn't explain something to a first-year undergraduate, then you didn't really understand it. Can everything in physics today be explained in such a simple manner?**
Since relativity and quantum mechanics came along you can kiss your imagination and intuition goodbye. The Higgs field is an example of something that you have a hard time explaining to anybody, and so is the fact that nothing can go faster than light. You need mathematics to cope.

What do you make of the huge public interest in the Higgs story?
Well, that's certainly none of my doing. As you saw with the faster-than-light-neutrino story, you begin in the centre of attention and enjoy it, then a little while later the thing goes down the drain and you've got egg on your face. That's the risk you run and it's the same with the Higgs. Although it's very unlikely that CERN is wrong, can you imagine the damage that would be done if they were?

So you're not a fan of CERN's Large Hadron Collider (LHC) collisions being likened to the Big Bang?
There will be nothing that comes out of the LHC that even vaguely helps us understand the Big Bang any better. It's outrageous baloney coming from cosmologists and astrophysicists who want to share some of the success of particle physicists. The Big Bang had an incredible amount of energy — a whole Universe-worth — whereas the amount of energy in the LHC collisions is nothing by comparison.

Do you ever discuss particle physics with your wife?
No. That's the loneliness of the theoretical physicist. I never was really capable of discussing physics with my family, with the exception of my daughter, who was also a particle theorist before moving into the banking world.

Who should get the Nobel prize for the Higgs?
There is only one guy who predicted that

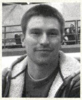
there should be a particle with non-zero mass: Peter Higgs. Robert Brout and François Englert discussed the same mechanism two months earlier, but they never got around to predicting this scalar particle. So Higgs should get the main credit, then Englert in second place. Brout is dead, so that leaves Tom Kibble, who I would say was working more on reality than Higgs was. Kibble is the most modest man on Earth and he deserves a lot more credit than he gets.

What keeps you awake at night?
There are still lots of problems that the Higgs doesn't solve. Why are the elementary particles arranged in three different families that are the same in all but their mass (for example, the electron, muon and tau)? You cannot build a machine to answer that type of question. But that is also the sad thing about the Higgs, which is why I wanted there to be no Higgs boson: it closes the last door on the standard model.

Will this end our subatomic exploration?
Oh no, this will never happen. Exploration is part of human activity. Suppose that we decided not to build any more particle colliders, and 50 years from now we're sitting there in a room staring at each other. Don't you think that at a certain point someone will say: shouldn't we at least have a look? ■

Interview by **Matthew Chalmers**, *a freelance science writer based in Bristol, UK.*

Question posed by Ragnar Stroberg, an experimental nuclear physicist studying for his PhD at Michigan State University.



R. STROBERG

I disagree with the notion that imagination and intuition are no longer viable in modern physics. If anything, they are more important than ever. At the core of both special and general relativity are brilliant examples of intuition and imagination. Einstein's thoughts of shooting a beam of light from a moving train paved the way for the mathematical development of his theories. That said, I admit I would have a hard time explaining the Higgs field to anyone — let alone a first-year undergraduate. But Feynman would say this just means I don't understand it well enough.

➲ SCIENTIFICAMERICAN.COM
You can visit Ragnar's 30 Under 30 page: go.nature.com/PAyDn7

ROLF SCHULTES

Q&A WILLIAM PHILLIPS

In search of ignorance

William Phillips is a physicist at the National Institute of Standards and Technology in Gaithersburg, Maryland. He was joint winner of the 1997 Nobel Prize in Physics, awarded for the development of laser cooling and trapping methods, and is still beguiled by the lure of the unknown.

What do you think are the challenges for young researchers these days — and how do they compare to when you were a graduate student?
It is always difficult to compare generations. When I started graduate school in the 1970s, it was right at the time when lasers were becoming a useful tool for atomic physics. That was an incredibly exciting time — almost everything you did would be something new. Although the laser was invented in the 1960s, it was only by about the 1970s that it was possible to buy a laser that you could tune to an atomic resonance, shine at atoms in a gas and excite them. This changed experimental physics in a very deep way.

If I look at physics as a whole, I am not sure that I can easily tell what gives people an advantage today, but I can see some opportunities. One of the things that happened during this year's Lindau meeting was the announcement from CERN of the discovery of the Higgs boson. For many people, it completes the standard model of particle physics, and there are two ways in which you can look at this. One is to say that it closes a chapter of physics; I think that is a very disappointing way of thinking of the discovery. A better way is to realize that it tells you where to go next. If you want to do something else, then you should look outside the standard model. There are lots of things that are not apparently included: dark matter and dark energy, for example (see 'Out of the darkness', page S2). These are areas where we can focus our attention.

We live in an incredibly exciting time for physics. We don't know what 96% of the Universe is made of — what can be more exciting than that? We have 4%, which is matter that we know about; 23% is dark matter, which we cannot yet identify; and the remaining 73% we know even less about. When I was a graduate student we didn't know what we didn't know about the Universe. Now we know what we don't know — and that is really a good place to be.

> ***"For a scientist, being somewhere where there is a great deal of ignorance is the best place to be."***

Lasers are used to trap and cool atoms to fractions of a degree above absolute zero, so that all atoms are in the same quantum energy state in a Bose–Einstein condensate.

PASCAL GOETGHELUCK/SCIENCE PHOTO LIBRARY

Did you ever suspect how far the field of cold atoms would go?

When I first started working on cold atoms in 1978, the motivation was to make a better atomic clock, which we achieved. It's amazing that all this work on cooling is useful for the thing we first imagined it would be. But there are so many other things — our inability to see where a field is going is an almost universal fact of the way we do research. When I was a postdoc I was studying atomic beams of sodium using tunable lasers and separately the Bose–Einstein condensation (atoms cooled to a very low temperature all sharing the lowest energy state) of atomic hydrogen. After finishing my postdoc, I first started to use lasers to slow atomic beams, so that the slow atoms could be used for atomic clocks. It was not in my mind that this was going to be a route to Bose–Einstein condensation. In fact, it turned out to be *the* route to a Bose–Einstein condensate (whereas the initial approach we took turned out not to be a particularly good way to make one). Finally, when we did have Bose–Einstein condensates, I don't think any of us imagined where this work would lead: into a whole new field of quantum degenerate gases (ultracold gases that exhibit quantum-mechanical behaviour).

There is an important lesson here: you must not imagine that you can see where a field is going, or think that you are smart enough to know where we should devote our resources if we want to accomplish a particular goal. Almost always, things turn out to be so much richer than you imagine. This is a lesson for everyone: for people who are going into research, and for people who are funding research. You just don't know.

Can we ever predict the future when it comes to science?

There are some subfields of science that become mature, full of things that have become well understood. But very often this is a prelude to finding out that there was something you were not paying attention to. All of a sudden you don't understand things you thought you understood. Atomic physics was like that in the 1950s — it was kind of boring. People thought that we understood quantum mechanics and hence everything about atoms. And then the laser came along, and we started doing experiments that nobody had even imagined. We go into the lab and learn new things every day.

What do you think is the most beautiful experiment in atomic physics today?

Let me choose two things. One is the idea of artificial fields. To take a neutral atom and make it behave as if it is a charged particle in a magnetic field — this is amazing. Second is the ability to image individual atoms and distinguish them from their neighbours. You can make a two-dimensional array of atoms and see where every one of those atoms is and how they have arranged themselves. These experiments open up opportunities for doing things that we could not even imagine a few years before.

> ***"You must not imagine that you can see where a field is going."***

✱ If you were a graduate student now, would you choose a different career?

I have seen a tremendous amount of excitement during my time, in particular the discovery of sub-Doppler cooling and of optical lattices. However, if you were to ask me what was the most exciting time in my whole career, I would say right now! There are so many incredible things going on: artificial fields and seeing individual atoms; recreating phenomena that occur in solids; 'atomtronics' — making circuits with atoms. This is why atomic physics and cold-atom physics is an exciting place to be.

At the same time, there is amazing work being done in other areas of science that would also interest me. For example, in genomics there are so many things that we still don't know. We don't know how most of our characteristics depend on the genome or on non-genomic things. And for a scientist, being somewhere where there is a great deal of ignorance is the best place to be.

Is it possible to be both a successful researcher and a good teacher and mentor?

Teaching and research are linked in an important way and are mutually beneficial. Unless somebody is actively engaged in research it is very difficult to teach at the level required to produce the next generation of researchers. People active in research are thinking about things in a cutting-edge way. But this works both ways — teaching helps you do research. That's because when you teach you are forced to re-examine the subject matter that you thought you knew well. You have to understand it in a way that is deeper than you ever understood it before. This gives you a much more mature understanding and raises questions in your own mind that will be important for your research. When you teach, students ask questions that you don't have answers for, and those questions make you think about new directions. That's why nobody who does research should ever be divorced from the mentoring of students. Those of us who supervise students and postdocs learn as much — if not more — from them as they learn from us. ■

Interview by **Iulia Georgescu**, *associate editor for* Nature Physics *in London.*

✱ Question posed by Martina Abb, a physicist at Southampton University investigating nanoscale optical devices.

M ABB

I really admire how enthusiastic Phillips is about his continuing research, even after having achieved so much. And I agree that there are an incredibly large number of interesting fields of research where there's still lots to be done. Genomics is definitely among them, although it's not the one I would choose if I were to change my field of research.

➲ SCIENTIFICAMERICAN.COM

You can visit Martina's 30 Under 30 page: go.nature.com/zsTPza

ROLF SCHULTES

Q&A ROY GLAUBER

From Los Alamos to Stockholm

Roy Glauber shared the Nobel Prize in Physics in 2005. He was cited for his contribution to the quantum theory of optical coherence, which explains why many experiments, using lasers for example, can only be understood if light is explicitly considered to be granular (composed of individual photons). His work laid the foundations of a field now known as quantum optics.

What sparked your interest in science?
My father was a travelling salesman and as soon as I was two years old, he took my mother and me on the road with him. We travelled all over the eastern half of the United States until I was about six years old. During this time, my mother — who had trained as an elementary school teacher but fell pregnant before she could start teaching — got me interested in all sorts of things, including science. My mother was my steady companion; she talked to me all the time and showed me many interesting places during our road trips.

I later started building model airplanes and, when I was around 12 years old (in 1937), I developed a passionate interest in astronomy. I read all the books in the local library on the subject and, in an encyclopedia, I found instructions on how to build a telescope using magnifying glasses. I built one, but the images were very poor and blurry, and they had rainbows around the edges. I quickly determined that a better idea was to build a reflecting telescope; in other words, I had learned very early on to look critically at what I had read and done.

During the Second World War you worked on developing the atomic bomb. How did that period influence you?
When I went to Los Alamos, I was 18 years of age and had been in college for just over two years. I was one of the youngest people working there — although the whole effort consisted of young people. The director of the project, Robert Oppenheimer, was only in his mid-thirties, and he was one of the oldest. There were not many older scientists who were able to move out to that very remote location.

I was working on the problem of neutron diffusion, and while my position was a junior one, I found that I was able to do the same sort of work as other far more educated and experienced scientists. That gave me the sense that I could hold my own among more senior people, and that I could be a scientist who contributed something. That's an uncertainty that many kids start out with — the sooner they overcome it, the sooner they can contribute.

> ***"Anyone undertaking scientific work imagining that they are working for the Nobel Prize is out of their mind."***

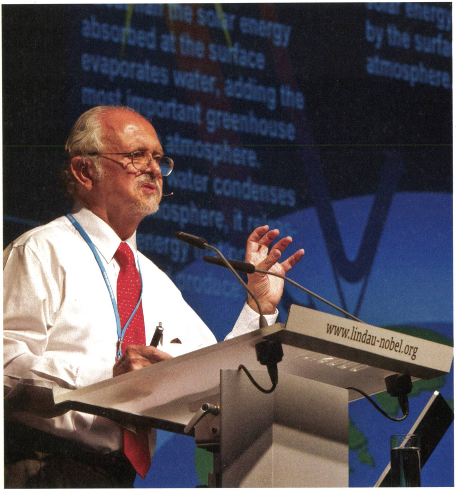
You were studying at Princeton at the same time that Albert Einstein was on the faculty. Did you have any interactions with him?
Einstein was far and away the most fascinating presence at the Institute for Advanced Studies in Princeton, where I spent a total of two years, from 1949 to 1951. But I had little to do with him. Einstein — already in his seventies — was not particularly active any more, and was critical of the developments that had taken place in quantum mechanics, which were upsetting his notion of causality. Despite being a pioneer of quantum theory, he had begun to have doubts about it. By the time I got to the Institute, he had really opted out of engagement.

Consequently he had nothing to do with people who were interested in quantum theory, as I was. That was very sad: we would have sacrificed a great deal to hear his criticisms or anything he had to say on the subject. But he wasn't available; he would sit in his office behind a closed door. I only encountered him when he was walking home. On one of these occasions I asked him whether I could take his photo, and he let me. I showed that photo for the first time in my Nobel lecture.

When you started working on quantum aspects of light — the work that would eventually earn you the Nobel Prize — did you have any idea that it would be so influential?
No. I thought it was a worthwhile pursuit but I did not see it as one of the great scientific questions of the age. It was a controversial topic at the time. But I was not thinking even remotely of the Nobel Prize, and I do not think that anyone else working in the field did either. After all, it took the Nobel committee more than 40 years to recognize the work. In fact, anyone undertaking scientific work imagining that they are working for the Nobel Prize is out of their mind. You should do what you know. If you think you have a fresh insight, you communicate it to the world and let the pieces fall where they may.

What is your advice to young people with an interest in science?
Find out what you think is really interesting and try to determine whether you have any skill at it. If you do, stick with it — the purpose is to generate new knowledge. I don't think that everyone should be a professional scientist, but many more kids should give it a try. ■

Interview by **Andreas Trabesinger**, *freelance science writer based in Brugg, Switzerland.*



CH. FLEMMING/LINDAU NOBEL LAUREATE MEETING

Q&A MARIO MOLINA
Socially responsible science

The first Mexican-born scientist to become a Nobel laureate in chemistry, Mario Molina shared the 1995 prize for his role in discovering the threat posed by chlorofluorocarbons (CFCs) to Earth's ozone layer. An optimist who passionately pursued science from a young age, Molina now focuses on finding practical solutions to environmental challenges.

At Lindau, you presented an optimistic outlook on our ability to tackle climate change, despite the scale of the challenge. What makes you so upbeat?

My optimism comes from the fact that society is quite capable of fixing problems, as it did by acting on the problem of CFCs. On climate change, the message from energy experts is that solutions are possible at a reasonable cost. Some of the solutions, such as becoming more energy efficient, make sense irrespective of the climate change problem and would be inexpensive. Others, such as capturing and storing carbon dioxide, are expensive. But the problem can be addressed at a relatively modest cost — in the order of 1–2% of GDP, which is smaller than the cost of the potential damage caused by climate change.

Why, then, has the climate change problem not been solved?

It is difficult for society to undergo the enormous changes required to address the problem. Furthermore, there are many powerful interest groups that have launched a well-financed public-relations campaign involving the media to discredit climate change science.

This year at Lindau, physicist Ivar Giaever questioned the role of human activity in climate change. How would you advise budding scientists to respond to climate change sceptics?

As a scientist, you cannot be an expert on everything. I would advise young scientists to be well informed on important societal issues that involve a scientific understanding of the problem, such as climate change. They should also be selective in their sources of information. With climate change, they have to be able to differentiate science from non-science — opinion from fact.

You have said that certain facts in science, such as the existence of molecules, are incontrovertible, but that uncertainties remain in climate science. What would it take to make human-induced climate change an incontrovertible fact?

We shouldn't be looking for certainty in climate change as it would take too long to make climate change incontrovertible. Instead we have to think of climate change as an issue of risk management. Society is constantly evaluating risk. As an analogy, at some point measurements of stratospheric ozone depletion were considered to point to the risk of a hole developing in the ozone layer. Society has to act based on such risks.

The basic conclusion of the scientific community is that the climate is indeed changing as a result of human activities, with potentially very serious consequences for society. The scientific community is aware that the current understanding of the science of climate change is not perfect and that much remains to be learned, but enough is known to estimate the probabilities that certain events will take place if society continues with 'business as usual' emissions of greenhouse gases.

What is the focus of your current work?

I hold a position at the University of California, San Diego, and also direct the Centro Mario Molina in Mexico City, where my focus is on science policy and communication. I collaborate with other experts, such as economists and engineers, to advise governments — including the Mexican government — on what makes sense in terms of sustainable development for greenhouse gas policy. For example, we are working on initiatives for Mexico such as providing incentives for solar energy to heat water, introducing fuel economy standards, and launching a pilot solar energy plant.

We are helping the government to consider regionally relevant policy options to tackle climate change, such as the use of photovoltaics in rural areas. We are also looking to other nations for inspiration: one very successful programme, for example, is the Trans Millennium rapid transit system in Bogota, Colombia. By dedicating one lane to buses, they provided a cheaper and complementary alternative to subways. We are learning from such

> ***"It's fine to be an advocate as long as you remain honest with the science."***

CARLOS CAZALIS/CORBIS

initiatives and are especially interested in policies with co-benefits. Measures that address greenhouse gas emissions, for example, can also improve air quality. Our overarching goal is to help achieve sustainable development, and climate change is part of this. There are other issues though: reducing congestion, for example, is important for improving the quality of life in cities.

When did you first become interested in science?

I became interested in science as a child after reading biographies of famous scientists. Later, when I was about ten, I received a chemistry set as a present from my parents. I still remember how thrilled I was the first time I used a toy microscope to observe a drop of dirty water, teeming with life. Years later I took control of a bathroom that no one used at home and turned it into a laboratory, and I carried out more sophisticated chemistry experiments with the help of an aunt who was a chemist.

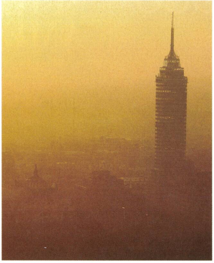
If you were starting out as a young climate researcher, where would you focus your efforts?

For chemists, many interesting questions remain. We still need a better understanding of how our planet functions and of atmospheric cycles, such as the nitrogen cycle. The science of extreme weather events is evolving rapidly and holds exciting possibilities for advancing understanding of climate change.

The work I undertook for my PhD was fundamental chemistry. We were able to measure concentrations of various compounds in the atmosphere and measure chemical reaction rates in the lab, and thereby confirm the composition of the atmosphere. Studying the urban atmosphere is much more complicated because it contains many more compounds and hence has many more reactions, not all of which can be well characterized in the lab. There is no shortage of interesting questions — the challenge is to find a way of addressing something that has relevance to society or to advance fundamental science in some way.

Would you advise young researchers to be as outspoken on the urgency of the climate problem as you were on the risk of stratospheric ozone depletion?

Yes. I think they should speak out to make sure that society moves in the right direction. I hope that speaking out in public is no longer the problem it was when we made our ozone discoveries. It was not well regarded at the time for scientists to talk to the media. It's fine to be an advocate as long as you remain honest with the science. Science can tell you the consequences of a given action, but our values tell us what to do with that information. Science cannot tell us to be better people.



Mexico City: one of the cities likely to benefit from the work of the Mario Molina Centre on sustainable development.

*** If you could give pre-university students one piece of advice, what would it be?**

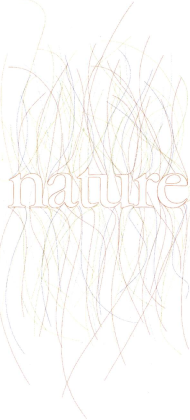
To be flexible. It's important to recognize that you might not actually like lab work or going to the mountains, so find out what you are good at and pursue it with passion. Also, I would advise them to be well informed and to start with a good foundation in science. You don't need to memorize the Periodic Table but you do need to understand how Mother Nature works! ■

Interview by **Olive Heffernan**, *freelance science writer and former chief editor of* Nature Climate Change *in London.*

*** Question posed by Minnie Mao, a postdoctoral researcher at the National Radio Astronomy Observatory in Socorro, New Mexico.**

The advice is good and I'll definitely be keeping it in mind during my career. But I find it interesting that the advice is so science-centric. I wonder how his answer would have changed had I asked more explicitly about work–life balance.

➲ SCIENTIFICAMERICAN.COM
You can visit Minnie's 30 Under 30 page:
go.nature.com/BQ4cKA



nature

A key thread in your scientific career

Whether in print, online or mobile, *Nature* is always at the forefront of science and is your forum to read, watch, listen and engage with key research, news and opinion.

How will you access the next landmark discovery?

➲ NATURE.COM/NATURE

nature publishing group npg

naturevideo
Confronting the Universe

in association with
SCIENTIFIC AMERICAN™

5 short films on physics

View the films on nature.com/lindau/2012*
and on iTunes and the *Nature Video* channel on YouTube

This summer, Nobel laureates and young physicists from around the world gathered in Germany for the annual Lindau Meeting. We filmed five debates on issues that matter to the current generation of researchers. Is dark matter real? How can we solve the looming energy crisis? How is physics perceived by the public? Watch the films to uncover the discussions and disagreements that emerged.

Access the Lindau Video site on your smartphone

*Published weekly from the 19th Sept – 10th Oct 2012
Image: NASA/ESA

Supported by:

MARS

MEETINGS OF NOBEL LAUREATES LINDAU®

nature publishing group npg

Mars, Incorporated

www.mars.com

Extending the possible

The 2012 Lindau Meeting provided a fitting venue for an early-morning discussion on the importance of collaboration in driving scientific discovery and extending the range of the possible.

Extending the possible through novel ideas and collaborative partnerships.
MARS incorporated
MEETINGS OF NOBEL LAUREATES LINDAU
CONTINUUM TECHNOLO

Dudley Herschbach (Centre Left), Amanda Peters Randles (Centre Right) and Chris Nagel (Right) discuss the role of cross-sector partnerships during the Mars-hosted panel discussion moderated by Adam Smith (Left).

"We believe that collaboration among the brightest scientists across all the sectors is essential to help address society's grand challenges," stated Harold Schmitz, Mars Chief Science Officer, as he introduced the Mars-hosted panel discussion at this year's Lindau Nobel Laureate Meeting.

Extensive collaboration between academic, government, non-government and industry sectors has long been central to Mars' approach to science. Over the years, Mars has been proud to collaborate on projects as diverse as mapping the cocoa genome, examining the role of genetics and nutrition in animal and human health, developing new approaches to food safety issues and studying the role of agroforestry in promoting biodiversity.

So it was fitting that this year's panel discussion should bring together scientists from across disciplines, sectors and generations to discuss the merits, and potential pitfalls, of a collaborative research approach. The discussion involved an audience of 150 young researchers and the panellists included Chemistry Nobel Laureate Dudley Herschbach, Chris Nagel, Founding Scientist of research company Continuum Energy Technologies, and Harvard University PhD student, Amanda Peters Randles.

Moderated by Adam Smith, Editorial Director of Nobel Media, the panellists shared examples from the world of physics to highlight how cross-sector partnerships can create a platform on which scientific breakthroughs can be appropriately debated, tested and validated to accelerate the development of solutions to many of society's grand challenges.

Dudley Herschbach, who won the Nobel Prize for Chemistry in 1986 for his contributions concerning the dynamics of chemical elementary processes, spoke of the myth of the "lonely genius" and emphasised that science is a shared enterprise: "All of science is collaborative – you can't do it by yourself." This point was supported by Amanda Peters Randles who spoke of the crucial role of collaboration in her research into cardiovascular disease and the challenges involved in managing an international team of doctors, physicists and computer scientists.

These contributions sparked questions and insights from the student audience, including topics such as when to bring in a collaborator, how to build a research profile and 'big' versus 'small' science. After the event, Dudley Herschbach spoke about this open exchange and noted: "Events such as this offer a unique opportunity for dialogue that not only crosses sectors and disciplines, but also moves across generations. By engaging with tomorrow's leading thinkers, we are equipping ourselves to deal with the challenges we face – both today, and in the future."

Mars' relationship with Lindau

Mars has been a key supporter of the Lindau Nobel Laureate Meetings since 2007. In 2008, Pamela Mars, member of the Mars family and board of directors, was appointed to the Honorary Senate of the Lindau Foundation in recognition of Mars' history of scientific achievement and commitment to scientific dialogue. Mars hosted a panel discussion at last year's Lindau meeting on 'Productive Partnerships' which featured the views of Laureate Elizabeth Blackburn and Frank Mars, President of Mars Symbioscience – the global health and life sciences segment of Mars. In 2010, the Mars-hosted panel discussion addressed the role of science in society and involved Laureate Françoise Barré-Sinoussi and Howard-Yana Shapiro, Global Staff Officer of Plant Science and External Research at Mars.

Uncommon collaboration

Mars research highlights

At Mars, we have a fundamental belief that evidence-based science is not only crucial to our ongoing success as a business, but also plays a key role in tackling a wide range of societal challenges, from climate change to food security. Yet we recognise that many of the challenges we face are too big to address alone and require the perspectives, skills and tools of diverse partners.

We aim to develop multi-disciplinary partnerships in which Mars scientists collaborate with researchers from world-renowned organisations to jointly advance scientific understanding. These uncommon collaborations are at the heart of our approach to science and over the years have resulted in a number of achievements.

Sequencing the cocoa genome

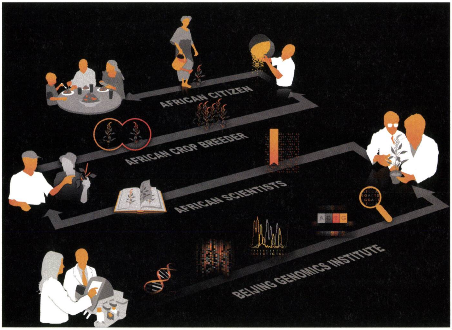
There are 6.5 million cocoa farmers globally and 40 million people whose livelihoods depend on the crop. Limited research resources in origin countries have left cocoa production facing major challenges, such as pests and disease, water shortages and deteriorating tree and soil condition.

In 2010, Mars, IBM and the US Department of Agriculture (USDA) completed a two-year project to sequence and annotate the cocoa genome. The sequencing project used a combination of the most modern DNA sequencers available, as well as three complementary bacterial artificial chromosome libraries from the highly homozygous cultivar Matina 1-6, which is considered by many experts to be the progenitor of at least 96% of all the cocoa cultivated in the world today. By analysing the DNA of hundreds of cocoa trees in South America, Mars' scientists were able to identify 10 distinct structure groups of the tree and their exact origins. To date, Mars research has identified more than 35,000 unique genes within the cocoa genome.

To allow scientists to apply this knowledge for the benefit of cocoa growers, the genome findings were released into the public domain through the Public Intellectual Property Resource for Agriculture and the Cacao Genome Database. Breeders across the world are already using this knowledge to identify traits of disease resistance, enhanced yields, water and nutrient-use efficiency and climate change adaptability among the world's cocoa. Ultimately, the results will lead to quicker, more efficient breeding and allow farmers to plant healthier, stronger, more productive cocoa that will improve their yields and income.

African Orphan Crops Consortium

Per capita food yields have been declining in Africa for decades and the Intergovernmental Panel on Climate Change (IPCC) reports that, by as early as 2020, global warming may reduce rain-fed crop yields in some African countries by as much as 50%. 96% of African agriculture is rain-fed rather than irrigated.



The African Orphan Crops Consortium brings together an international team of experts with the aim of driving advances in plant breeding that will have a direct impact on the lives of 600 million Africans involved in farming.

To address these challenges, the African Orphan Crops Consortium (AOCC) was launched at the Clinton Global Initiative meeting in 2011. Mars is a lead partner of the AOCC, which brings together experts such as the African Union, The New Partnership for Africa's Development, the Beijing Genomics Institute, World Wildlife Fund, Life Technologies, University of California, Davis and many others. Working together, the Consortium aims to genetically sequence and release into the public domain the genomes of 100 of Africa's most important food crops to help breeders make them more high-yielding, nutritious and resilient. These "orphan crops" are crucial to African farming families, but have been largely ignored by science because they are not traded internationally.

To ensure this information reaches farmers, the AOCC will establish the African Plant Breeding Academy in Ghana and Kenya. The academy will train 250 plant breeding scientists and 500 technicians over five years. The goal is to educate African plant breeders in the application of genomic information to crop improvement. This will accelerate the rate of genetic improvement to increase the yield and nutritional quality of African staple crops. Taken together, the Consortium's work can help Africa tackle food scarcity and malnutrition, while improving the incomes of the 600 million Africans involved in farming.

Rice sustainability

Carbon footprint is driven by three greenhouse gases (GHG): carbon dioxide, methane and nitrous oxide. Unlike most crop systems, rice produces both nitrous oxide and methane and has one of the highest impacts of any crop in terms of water use and the release of methane.

Mars' strategy on rice sustainability is led by David Mackill, former principal scientist of the International Rice Research Institute (IRRI) and a leading expert in rice plant breeding, genetics and biotechnology. Mackill's previous work has been based on the use of precision-breeding techniques to develop a high-yield, flood-tolerant rice that can withstand up to 20 days under water. Mackill's focus at Mars is on establishing a global network of rice scientists, research institutions, and government agencies to generate groundbreaking new research to support sustainable rice production and breeding programmes.

In addition, Mars is currently funding a programme with the University of California, Davis and the University of Arkansas to establish base-line GHG emission and global warming potential (GWP) data for rice systems in Europe and the US and identify mitigation strategies that lower the GWP of rice systems. The ultimate aim is to develop robust and reliable recommendations that are both universally applicable and can be adjusted to a regional farmer's perspective.

Early results indicate that the lowest GHG emissions per unit of grain yield are achieved using good management practices and just enough nitrogen fertilizer to achieve optimal grain yields. This is good news for both farmers and the environment. Preliminary results from these studies suggest that dry seeding or intermittent irrigation can reduce the length of time fields are flooded and thus cut GHG emissions. Rice varieties may affect levels of GHG emissions, which suggests that future efforts could be directed toward selecting rice varieties that emit less GHG.

Cocoa flavanols

A growing body of scientific evidence suggests that dietary flavanols — a phytonutrient found in cocoa, tea, grapes, red wine, apples and other fruit — helps support healthy circulation and cardiovascular health. Recent findings also provide compelling new evidence that regular consumption of dietary cocoa flavanols may improve cognitive function in elderly subjects with early memory decline.

Mars initiated its flavanol research programme in the late 1980s and today works with some of the world's leading research institutions in the field, including Columbia University, New York; the Salk Institute in La Jolla; the University of California, Davis; the University of California, San Francisco; the University of Reading, UK; the Heinrich Heine University, Düsseldorf, Germany; the Baker IDI Heart & Diabetes Institute, Australia; the European FLAVIOLA research consortium, and many other research groups around the world.

Mars adopts a holistic approach to cocoa sustainability that focuses on plant genetics, breeding, pest and disease control and post-harvest practices.

Mars' main focus areas to date have been in flavanol analytics and chemistry, flavanol preserving food processing technologies, and the investigation of the biomedical properties of flavanols. Mars' ongoing fundamental research centred on linking chemistry with human health has enabled significant scientific advancements, and may one day lead to meaningful dietary recommendations for cocoa flavanol intake. As part of its collaborative research programme, Mars engages in extensive dialogue with the scientific community and is committed to the peer-review process. Taken together, Mars' work is represented by the publication of over 140 scientific papers and approximately 100 patents. This includes our proprietary, patented Cocoapro® process that helps to retain the flavanols found naturally inside the cocoa bean, which are usually destroyed during normal processing.

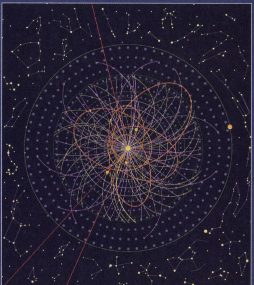
Further reading

- Schmitz, H.H. Science and society: what is the role and responsibility of corporations in this context? *Nature* Vol. 467, No. 7317_supp ppS1-S23 (2010). www.nature.com/nature/outlook/masterclass/pdf/mars.pdf
- Cacao Genome Database, *The Genomics, Genetics and Breeding Resource for Cacao Improvement*, www.cacaogenomedb.org (2012).
- Mars, Incorporated, *Sustainable Cocoa Initiative – Securing Cocoa's Future*, www.mars.com/global/brands/cocoa-sustainability-home.aspx (2012).
- University of California, Davis, *$40 million project to revitalize Africa's orphaned crops announced*, http://news.ucdavis.edu/search/news_detail.lasso?id=10021 (2011).
- Schroeter, H. *et al.* Recommending flavanols and procyanidins for cardiovascular health: current knowledge and future needs. *Molecular Aspects of Medicine* 31, 546–557 (2010).
- Mars, Incorporated, *Healthycocoa.com Library*, www.healthycocoa.com (2012); Mars, Incorporated, *Centre for Cocoa Health Science*, www.marscocoascience.com (2012).

Mars, Incorporated

In 1911, Frank C. Mars made the first Mars candies in his Tacoma, Washington kitchen and established Mars' roots as a confectionery company. Today, Mars has net sales of more than $30 billion and six business segments including Petcare, Chocolate, Wrigley, Food, Drinks and Symbioscience. As a private, family-owned company, we have the opportunity and freedom to take a long-term approach to science. In addition to our investments in product innovation, we fund and conduct fundamental and applied scientific research in areas where our business interests and society's interests intersect.

Among other activities, we are using our in-house scientific knowledge and our cross-sector research partnerships to boost supplies of sustainable raw materials, develop strategies for making our operations more sustainable, and improve the nutritional profile of our products.



nature

http://www.nature.com/nature/outlook/masterclass_2012

naturejobs.com

Upload your CV at *Naturejobs* and let's take your career sky high

Ready to take your job search to the next level? Upload your CV and cover letter at the new improved naturejobs.com and take full advantage of the world's largest science jobs board. Your saved information will be immediately available, so you can quickly and easily apply for one of the **10,000+ science vacancies** found at naturejobs.com.

To find out more, visit:
www.naturejobs.com/cvupload

Follow us on:

nature publishing group npg

Super capabilities are now within your reach
96

The new LightCycler® 96 Real-Time PCR System delivers the speed, analytical power and publication-ready results that make you a research superhero.

LightCycler®

Find out more about availability and features on www.lightcycler96.com

For life science research only.
Not for use in diagnostic procedures.

LIGHTCYCLER is a trademark of Roche.
NOTICE: This product may be subject to certain use restrictions. Before using this product please refer to the Online Technical Support page (http://technical-support.roche.com) and search under the product number or the product name, whether this product is subject to a license disclaimer containing use restrictions.

Roche Diagnostics GmbH
Sandhofer Straße 116
68305 Mannheim, Germany

© 2012 Roche Diagnostics.
All rights reserved.

Roche

RESEARCH

NEW ONLINE

Papers published this week at Nature.com/nature

PERSPECTIVE Why trial results in animal models are so hard to repeat in the clinic **p.187**

REVIEW As production methods improve, graphene will soon find new applications **p.192**

LETTER If crop yields reach their theoretical maxima, hunger could be history **p.254**

PUBLISHED 7 OCTOBER 2012

STEM CELLS

Organ-specific stem-cell boost

For stem cells to be of use in regenerative medicine they have to be available in quantity. Douglas Melton and colleagues now amplify pancreatic progenitor cells by co-culturing them with organ-matched mesenchyme, which acts as a self-renewal signal. They achieve a 65-million-fold expansion of human endodermal progenitors, retaining full developmental potential. This work paves the way for the efficient production of important cell types, such as insulin-secreting β cells, for therapeutic use.

Self-renewal of embryonic-stem-cell-derived progenitors by organ-matched mesenchyme
J B Sneddon *et al.*
doi:10.1038/nature11463

BIOPHYSICS

Excited-state RNA capture

The sequence and structure of RNA have long been recognized as being important for the molecule's biological roles. More recently, conformational dynamics have emerged as powerful regulators of RNA function. Elizabeth Dethoff *et al.* now describe an approach based on nuclear magnetic resonance that can capture previously inaccessible, highly transient, lowly populated excited states in RNA. The localized rearrangements in base pairing giving rise to these states affect function by changing the exposure of residues required for a specific biological process.

Visualizing transient low-populated structures of RNA
E A Dethoff *et al.*
doi:10.1038/nature11498

SYNTHETIC BIOLOGY

Complex synthetic gene circuitry

Synthetic genetic circuits tend to interfere with one another, a complication that restricts the number of circuits that can be used to program a cell. Chris Voigt and colleagues have mined a collection of DNA-binding proteins that depend on specific chaperone proteins to activate the transcription of their target genes, and combined them into complex, layered circuits. The authors have thereby constructed one of the largest genetic programs built to date, consisting of seven integrated sensors/circuits and eleven regulatory proteins.

Genetic programs constructed from layered logic gates in single cells
T S Moon et al.
doi:10.1038/nature11516

MOLECULAR BIOLOGY

The spatial architecture of protein function and adaptation
R N McLaughlin Jr *et al.*
doi:10.1038/nature11500

NEUROSCIENCE

Intrinsically determined cell death of developing cortical interneurons
D G Southwell *et al.*
doi:10.1038/nature11523

PUBLISHED 10 OCTOBER 2012

EARTH SCIENCE

Xenon depletion by magma oceans

The noble gas xenon is strongly depleted relative to argon in the atmospheres of Earth and Mars, when compared with the abundances in chondritic meteorites, which are generally considered as indicators of early Solar System geochemistry. Svyatoslav Shcheka and Hans Keppler offer a possible explanation for this apparent anomaly. They show that more than 1 wt% of argon can be dissolved in the abundant lower-mantle mineral $MgSiO_3$ perovskite, but that the solubility of xenon in this mineral is orders of magnitude lower. The crystallization of $MgSiO_3$ perovskite from a magma ocean in the very early stages of Earth history could therefore have concentrated argon in the lower mantle, with xenon being lost along with early Earth's atmosphere. The argon from the lower mantle was then gradually released into Earth's atmosphere, explaining its abundance relative to xenon.

The origin of the terrestrial noble-gas signature
S S Shcheka & H Keppler
doi:10.1038/nature11506

STEM CELLS

Primed for thyroid production

Sabine Costagliola and colleagues report a protocol that converts mouse embryonic stem cells into functional thyroid follicles *in vitro*. Overexpression of the transcription factors NKX2-1 and PAX8 directs differentiation towards thyroid follicular cells, which self assemble when treated with thyrotropin. The resultant three-dimensional thyroid follicles show hallmarks of thyroid function *in vitro*, and can rescue multiple symptoms when transplanted into athyroid mice. This work not only adds to our understanding of the molecular mechanism behind thyroid development, but also paves the way for regenerative medicine to treat congenital hypothyroidism, the most common congenital endocrine disease in humans.

Generation of functional thyroid from embryonic stem cells
F Antonica *et al.*
doi:10.1038/nature11525

PHYSIOLOGY

CaMKII determines mitochondrial stress responses in heart
M-I A Joiner *et al.*
doi:10.1038/nature11444

MATERIALS SCIENCE

Bonding and structure of a reconstructed (001) surface of $SrTiO_3$ from TEM
G-z Zhu *et al.*
doi:10.1038/nature11563

CANCER

Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation
J Landsberg *et al.*
doi:10.1038/nature11538

Tumours switch to resist
A Ribas & P C Tumeh
doi:10.1038/nature11489

EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations
M T McCabe *et al.*
doi:10.1038/nature11606

STRUCTURAL BIOLOGY

Structure of the agonist-bound neurotensin receptor
J F White *et al.*
doi:10.1038/nature11558

Snapshot of an activated peptide receptor
F Hausch & F Holsboer
doi:10.1038/nature11634

MR1 presents microbial vitamin B metabolites to MAIT cells
L Kjer-Nielsen *et al.*
doi:10.1038/nature11605

NEUROSCIENCE

Promiscuous vesicles

The unexpected finding that neurons can co-release two neurotransmitter molecules, dopamine and GABA, through a common mechanism provides a further advance in our understanding of the nervous system. See Letter p.262

JOHN T. WILLIAMS

The striatum, a part of the brain that regulates motivation, reward responses, feeding and movement, integrates signals from other brain regions such as the substantia nigra and the ventral tegmental area. The inputs from these two regions arrive through nerve fibres that release dopamine, a neurotransmitter molecule that modulates neuronal activity by means of a slow-acting process. However, on page 262 of this issue, Tritsch *et al.*[1] report that these modulatory neurons can also induce a rapid, short-lasting inhibition of striatal neurons by releasing another neurotransmitter, γ-aminobutyric acid (GABA). Surprisingly, the authors found that this GABA-mediated effect was dependent on the protein VMAT2, which is required for dopamine secretion, rather than on VGAT, a protein that was thought to be needed for GABA release.

Proteins such as VMAT2 and VGAT are called vesicular transporters, because they pump neurotransmitters from the neuronal cytoplasm into vesicles that are then unloaded at the synaptic junction between a sender neuron and a receiver neuron (Fig. 1). The released neurotransmitters bind to specific receptor proteins on the surface of the receiver cell, leading to changes in the cell's activity — by, for example, altering the flow of ions that cross the cell membrane through proteins called ion channels. Neurotransmitters can also bind to receptors on the sender cell (autoreceptors) that typically modulate the release or synthesis of the same neurotransmitter.

Dopamine-releasing (dopaminergic) neurons that project from the substantia nigra and the ventral tegmental area (VTA) form a heterogeneous population; these cells differ in the ion channels that they express on their surface[2], in their receptor proteins (such as G-protein-coupled receptors[3]) and in their vesicular transporters[4]. Some of the neurons can secrete glutamate[5–7] (a neurotransmitter and a precursor in the synthesis of GABA), and dopamine and glutamate are stored and released from the same synaptic vesicles in a subset of dopaminergic neurons located mainly in the VTA. Glutamate transport into vesicles has been shown to be dependent on a specific vesicular transporter (VGLUT2)[4].

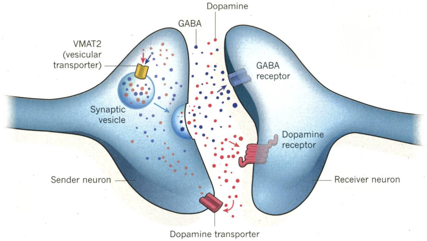


Figure 1 | Two messages in one parcel. Information can flow from one brain area to another — for example, from the substantia nigra to the striatum — through neural fibres that end in synaptic connections between neurons. In the sender neuron, various proteins (vesicular transporters) pack neurotransmitter molecules such as dopamine into vesicles and, when the cell is activated, the synaptic vesicles are discharged into the space between the two neurons. The neurotransmitters then bind to receptor proteins on the surface of the second neuron, triggering changes in the cell's activity. They can also be returned to the sender cell's cytoplasm by specific transporter proteins. Tritsch *et al.*[1] report that dopamine-releasing neurons projecting from the substantia nigra can use the vesicular transporter VMAT2 to store dopamine together with another neurotransmitter, γ-aminobutyric acid (GABA), in the same vesicles. As these compounds bind to receptors that have different effects on the receiver cell, their co-release allows dopamine-releasing neurons to modulate the activities of striatal neurons in various ways.

To clarify how dopaminergic neurons in the substantia nigra affect striatal activity, Tritsch *et al.* used an optogenetic technique that allowed them to turn the neurons on and off in brain slices obtained from genetically engineered mice. In this experimental set-up, brief pulses of light induced the neurons to secrete dopamine. Surprisingly, the authors found that this selective activation of dopaminergic neurons produced large, GABA-dependent inhibitory currents and small, glutamate-dependent excitatory currents in certain striatal neurons. The GABA-dependent currents were large enough to inhibit the firing of the striatal neurons. Even more surprising was the finding that knocking out VGAT (the vesicular GABA transporter) did not block GABA's inhibitory effect, raising the question of where the released GABA was coming from. The researchers went on to show that GABA's inhibitory currents were blocked by inhibiting VMAT2, suggesting that GABA is concentrated with dopamine in the same synaptic vesicles.

Further support for this conclusion came from a remarkable series of experiments. The authors found, for example, that decreasing the release of dopamine by inhibiting its synthesis, or by knocking out the expression of VGLUT2 or VGAT, did not affect the inhibition by GABA. By contrast, VMAT inhibitors blocked the secretion of both dopamine and GABA. Taken together, these assays indicate that VMAT2 alone is responsible for the accumulation of GABA in dopamine-containing vesicles. How GABA is synthesized in these neurons remains to be determined, because only a small subset of dopaminergic neurons seems to express the gene *Gad65* (also known as *Gad2*)[8], which encodes a key enzyme involved in the synthesis of the neurotransmitter.

Tritsch and colleagues' work shows, therefore, that co-released dopamine and GABA modulate the activity of the striatal output neurons in temporally and mechanistically distinct ways. It also suggests that, in the striatum, GABA-mediated inhibition dominates over glutamate-mediated activation. Moreover, the authors' results indicate that GABA is probably co-released by other dopamine neurons, and so it will be important to determine the physiological impact of this process in the striatum and in other brain areas *in vivo*.

Although the optogenetic technique was instrumental in identifying this process, the method is known to produce synchronous stimulation of many neurons, and so it is not clear whether GABA-dependent inhibition will be observed in the absence of this massive stimulation. It is worth noting, nevertheless, that dopaminergic neurons in subregions of the substantia nigra and the VTA tend to burst in relative synchrony as a result of a common input[9]. In addition, the fibres of neurons in these subregions tend to project to the same brain areas[2], and so it is possible that a phasic (intermittent) GABA-dependent inhibition occurs in selective areas.

As the authors point out, VMAT is expressed not only by dopaminergic neurons but also by neurons that release other monoamines (neurotransmitters that are synthesized from aromatic amino acids), such as serotonin, adrenaline and noradrenaline. These neurons would therefore also be expected to release GABA. If true, then this property can be added to the many similarities among monoaminergic neurons — for example, in their firing rate and pattern, in the co-release of glutamate and monoamines, and in the expression of their respective inhibitory autoreceptors. Transmission from monoaminergic neurons can thus induce not only slow, long-lasting modulation but also brief, phasic inhibition. In my opinion, this finding will change the way researchers think about the roles of monoaminergic neurons in the functioning of the brain. ■

John T. Williams *is at the Vollum Institute, Oregon Health and Sciences University, Portland, Oregon 97239, USA.*
e-mail: williamj@ohsu.edu

1. Tritsch, N. X., Ding, J. B. & Sabatini, B. L. *Nature* **490,** 262–266 (2012).
2. Lammel, S. *et al. Neuron* **57,** 760–773 (2008).
3. Ford, C. P., Mark, G. P. & Williams, J. T. *J. Neurosci.* **26,** 2788–2797 (2006).
4. Hnasko, T. S. *et al. Neuron* **65,** 643–656 (2010).
5. Sulzer, D. *et al. J. Neurosci.* **18,** 4588–4602 (1998).
6. Stuber, G. D., Hnasko, T. S., Britt, J. P., Edwards, R. H. & Bonci, A. *J. Neurosci.* **30,** 8229–8233 (2010).
7. Tecuapetla, F. *et al. J. Neurosci.* **30,** 7105–7110 (2010).
8. González-Hernández, T., Barroso-Chinea, P., Acevedo, A., Salido, E. & Rodríguez, M. *Eur. J. Neurosci.* **13,** 57–67 (2001).
9. Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A. & Uchida, N. *Neuron* **74,** 858–873 (2012).

ORGANIC CHEMISTRY

Single molecules put a ring on it

A variant of a classical reaction has been used to generate short-lived chemical species called arynes, allowing the one-step synthesis of structurally complex benzene derivatives from simple precursors. SEE ARTICLE P.208

JOHN T. S. YEOMAN & SARAH E. REISMAN

The reactions of arenes — benzene and its derivatives — have long been exploited by organic chemists for practical applications ranging from the production of simple commodity chemicals, such as polymers and dyes, to the synthesis of pharmaceuticals and other structurally complex compounds. One approach to the synthesis of arenes involves the generation of benzyne, a highly reactive, short-lived intermediate that can undergo a variety of chemical transformations. On page 208 of this issue, Hoye and co-workers[1] describe a clever strategy for the generation of arynes (benzyne-containing intermediates) that involves the heat-induced isomerization of a starting material, without the need for additional reagents. This transformation is complementary to existing methods for aryne generation, and could expand the synthetic utility of these well-studied intermediates.

Arynes have been of theoretical and practical interest to chemists for the past century. However, it was not until 1953 that an isotopic-labelling study conducted by the US chemist Jack Roberts and co-workers[2] provided the first compelling experimental evidence for the intermediacy of benzyne in a reaction. Organic chemists soon recognized the synthetic utility of arynes, and considerable effort has been devoted to the development of methods to generate these highly reactive species[3].

Classically, arynes are generated from haloarenes — arenes in which one or more hydrogen atoms have been replaced by halogen atoms — using either strong bases or reactive organometallic reagents (Fig. 1). Although these methods work well for simple haloarenes, the use of a strong base (or a reactive organometallic reagent) can lead to unwanted side reactions when arenes containing certain chemical groups are employed. Alternatively, arynes can be generated under 'milder' conditions that are compatible with a wider range of groups by oxidizing compounds known as aminobenzotriazoles, or by heating arenediazonium carboxylate salts (Fig. 1). However, the former substrates can be difficult to synthesize, whereas the latter can explode when handled.

A breakthrough for the practical application of arynes was the discovery[4] that exposure of compounds known as aryl silyl triflates to fluoride salts produces arynes at room temperature under nearly neutral conditions (Fig. 1). This mild reaction enables the use of complex aryne precursors and reaction partners, and has inspired a renaissance in aryne chemistry[5].

The strategies highlighted above can all be considered elimination reactions, in which a precursor arene loses one or more groups to form an aryne. Hoye and colleagues' approach, however, depends on a classic reaction known as the Diels–Alder cycloaddition[6], and therefore represents a fundamentally different strategy. The Diels–Alder cycloaddition is the most

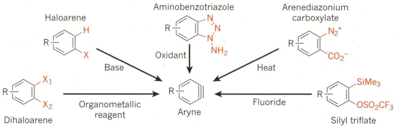


Figure 1 | Methods for preparing arynes. Arynes are synthetically useful chemical intermediates that can be prepared from a variety of starting materials using the general reactions shown. The groups highlighted in red are eliminated to provide arynes when subjected to the reagents or conditions indicated. In each case, the starting material contains a benzene ring. R represents any chemical group; X, X_1 and X_2 can be any halogen atom; Me is a methyl group. The chemical bond extending from the centre of each benzene (or aryne) ring indicates that the bond could be attached to any of the carbon atoms on the ring that is not depicted as having a chemical group attached.

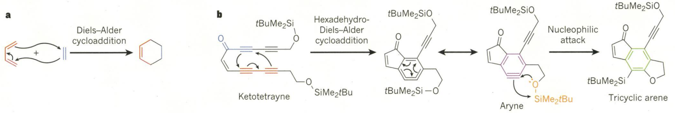


Figure 2 | A fresh twist on the Diels–Alder reaction. **a,** In the classic Diels–Alder cycloaddition reaction, a diene (red) reacts with a dienophile (blue) to form a cyclic product. **b,** While attempting to prepare a ketotetrayne, Hoye *et al.*[1] instead isolated a tricyclic arene. They propose that a diyne group (red) in the ketotetrayne undergoes a hexadehydro-Diels–Alder reaction with a diynophile group (blue) in the same molecule. The resulting compound can be thought of as a hybrid of two 'resonance' forms (double-headed arrow indicates resonance), one of which contains an aryne (purple). A pendant silyl ether group (orange) then reacts with the aryne to form the tricyclic arene, in which the aryne has become a benzene ring (green). Curly arrows indicate electron movement, and the dots on the oxygen atom in the silyl ether represent a pair of electrons. Me, methyl; *t*Bu is a tertiary butyl group, CMe_3.

widely studied and used method for synthesizing compounds that contain rings of six atoms[7] (Fig. 2a). It involves the simultaneous formation of two carbon–carbon bonds between a diene (a molecule that contains four carbon atoms and two carbon–carbon double bonds) and a dienophile (which has two atoms and one double bond) to make a six-membered ring. The reaction works for a wide range of dienes and dienophiles, proceeds with predictable isomeric selectivity, and is highly atom-economical: all of the atoms present in the reactants appear in the product, minimizing waste.

As is the case for many scientific advances, Hoye and colleagues' discovery of their 'hexadehydro-Diels–Alder' (HDDA) reaction occurred serendipitously, during their efforts to prepare a compound known as a ketotetrayne. Instead of the ketotetrayne, they obtained a tricyclic arene as the major product (Fig. 2b). They recognized that a cycloaddition process was taking place: two carbon–carbon triple bonds (alkynes) were reacting with a third alkyne triple bond from the same molecule to generate an aryne fused to one other ring. In terminology analogous to the diene and dienophile of the classical Diels–Alder reaction, the first two alkynes comprise a diyne group, whereas the third alkyne reacts as a diynophile. The transiently formed aryne then reacted with a fortuitously positioned silyl ether group in the molecule to yield the tricyclic arene as the final product.

Realizing the potential power of this new reaction to generate arynes, Hoye *et al.* investigated the scope and limitations of the transformation. The reaction tolerates substrates containing a variety of chemical groups, demonstrating the generality of the reagent- and by-product-free conditions for aryne formation. The authors found that electron-withdrawing groups on the diynophile accelerate the reaction, an effect that is characteristic of standard Diels–Alder cycloadditions. In addition to silyl ethers, the aryne intermediates can react with a variety of other chemical groups, such as alcohols (which contain OH groups), olefins (which contain carbon–carbon double bonds), amides (which contain NH groups), bromide ions (Br^-) and aromatic rings.

Hoye and colleagues' reactions are generally high yielding, and they provide rapid access to complex arenes that could be challenging to prepare by more conventional means. Although the HDDA reaction has been the subject of previous theoretical and experimental studies[8,9], the authors' report is the first to demonstrate practical applications of the transiently generated aryne intermediates and to thoroughly explore the substrate scope of these transformations.

The diyne and diynophile in the described HDDA reactions are contained within the same molecule, but one can imagine that a reaction between a diyne and diynophile from two different molecules would be an even more powerful method to construct structurally complex arenes. The prospect of such a reaction, and of myriad new aryne-reaction modes that will be enabled by the reagent-free conditions, will undoubtedly encourage continued interest in and investigation of aryne chemistry for years to come. ■

John T. S. Yeoman *and* **Sarah E. Reisman** *are in the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA.*
e-mail: reisman@caltech.edu

1. Hoye, T. R., Baire, B., Niu, D. Willoughby, P. H. & Woods, B. P. *Nature* **490,** 208–212 (2012).
2. Roberts, J. D. *et al. J. Am. Chem. Soc.* **75,** 3290–3291 (1953).
3. Kitamura, T. *Aust. J. Chem.* **63,** 987–1001 (2010).
4. Himeshima, Y., Sonoda, T. & Kobayashi, H. *Chem. Lett.* **12,** 1211–1214 (1983).
5. Tadross, P. M. & Stoltz, B. M. *Chem. Rev.* **112,** 3550–3577 (2012).
6. Diels, O. & Alder, K. *J. Liebigs Ann. Chem.* **460,** 98–122 (1928).
7. Nicolaou, K. C.*et al. Angew. Chem. Int. Edn* **41,** 1668–1698 (2002).
8. Cahill, K. J., Ajaz, A. & Johnson, R. P. *Aust. J. Chem.* **63,** 1007–1012 (2010).
9. Kawano, T., Inai, H., Miyawaki, K. & Ueda, I. *Tetrahedron Lett.* **46,** 1233–1236 (2005).

PALAEONTOLOGY

Cambrian nervous wrecks

Fossilized remains of an arthropod from the Cambrian period provide an unusual example of preservation of the brain and nervous system, and shed new light on when and how these tissues evolved. SEE LETTER P.258

GRAHAM E. BUDD

Even to palaeontologists, the fossil record can resemble the chaotic attic of an eccentric relative, stacked with ancient bric-a-brac of dubious usefulness. But the record has recently been throwing up some surprises that are bringing new order to this jumble. Our concept of dinosaurs, for example, has evolved from what were essentially bolted-together lumps of bone into living creatures covered in graceful feathers — and in colour too[1]. Other fossil finds have brought changes to the scale of our understanding of evolution. For example, the discovery of exceptionally well-preserved fossil muscle fibres throughout the record[2] and fossilized embryos from at least the Cambrian period[3], some 500 million years ago, have provided remarkable insight into the fine-scale evolution of these tissues and life stages. Now, on page 258 of this issue, Ma and colleagues[4] describe preserved nervous tissue from the

Cambrian — a find that grants palaeontologists access to the exclusive zoological club of those who study the brain and nervous system.

Certain tissue types, such as bone and shell, are much more likely to be preserved in the fossil record than others, owing to their mineralized and thus relatively refractory nature. Soft tissues, such as muscle or even the cuticle of arthropods (joint-legged invertebrate animals), which is largely proteinaceous and therefore liable to decay, are less likely to withstand oxidation, scavenging, mechanical damage and all the other vicissitudes that conspire to make the work of a palaeontologist so testing. Tissues also preserve differently under different conditions — 'bog people', for example, are often all leather and no bones because of the acidity of the peat they were preserved in. This variable survival is captured by the idea of preservation potential, and understanding this concept is critical when trying to reconstruct ancient organisms, because it places important limitations on what a difficult-to-interpret structure in a fossil might be. It is highly unlikely, for example, that delicate blood vessels are going to leave a trace in a fossil found in sandstone deposited in a high-energy environment.

The concept of preservation potential has recently been receiving empirical support from decay experiments — a useful, if rather smelly, approach to understanding the fossil record[5]. These studies have taught us that descriptions of fossils that portray kidneys and gonads rightfully raise eyebrows. Yet, somewhat surprisingly, they also suggest that the brain and nervous system might have a decent chance of preservation, at least in vertebrate animals[5]. The situation in invertebrates seems less clear, although possible examples have been reported[6,7]. Ma *et al.*, however, have now provided convincing evidence for the preservation of the brain and partial nervous system in fossils of a Cambrian arthropod, and one of the most controversial and interesting arthropod species to boot, *Fuxianhuia*.

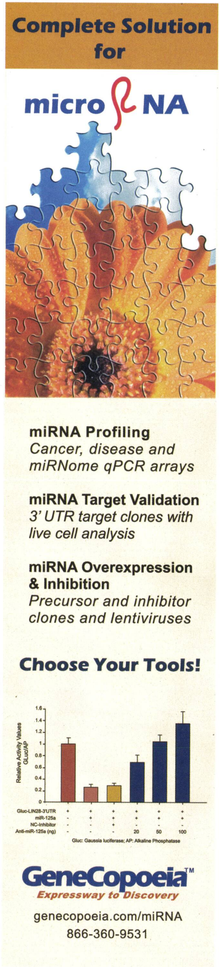
The taxonomic position of *Fuxianhuia* has been the topic of some dispute, but it is generally considered to lie close to the most recent common ancestor of the living arthropods[8]. As such, this fossilized material is ideally placed to contribute to neurophylogeny, the relatively new exploration of brain and nervous-system architecture in a phylogenetic context[9]. What does the most informative of these *Fuxianhuia* specimens tell us? Starting at the top, the fossil displays two large, faceted eyes on stalks (see Fig. 1 of the paper). Within the stalks, a dark, iron-rich material traces what Ma and colleagues interpret to be three neuropils (concentrations of nervous tissue) from which optic nerves lead towards the brain. Here, the preserved material is divided into three regions that correspond to the classical arthropod brain ganglia of the protocerebrum, deutocerebrum and tritocerebrum (a ganglion is a mass of nerve cells). The authors suggest that all three regions lie, at least in part, in front of the stomodeum — the part of the gut that leads directly to the mouth. From the deutocerebrum and tritocerebrum, stout nerves lead to the antenna and to what is presumably another appendage, respectively. This latter structure has proved problematic to describe in other *Fuxianhuia* fossils, and its very existence has sometimes been questioned[8].

This fossil provides the most convincing, and certainly the oldest, description of nervous-system tissue in a fossil arthropod. So what do these remarkable results imply? The most striking feature, as the authors themselves stress, is how similar the tripartite brain of *Fuxianhuia* is to that of most modern arthropods, including crustaceans and insects. The exception to this are the branchiopods — the best-known example of which is *Daphnia*, the elegant water flea — which have a considerably simpler brain, in that they lack both complex optic neuropils and a prostomodeal tritocerebrum. If *Fuxianhuia* belongs in the stem group of all arthropods, then it follows that the organism's complex brain organization evolved relatively early within this phylum. If this is the case, the authors suggest, then the predecessors of extant arthropods with less complex brains (which includes not only the branchiopods, but also spiders, scorpions and their relatives) must have at some stage simplified their arrangements.

However, there are two potential alternatives to this far-reaching conclusion. It is possible that the arrangement in *Fuxianhuia* is convergent to that in the modern crustaceans or insects; in other words, similar brain assemblies to that reported for *Fuxianhuia* evolved again in later arthropods. Or it may be that we need to rethink the systematic position of *Fuxianhuia*. That latter option would entail a substantial rearrangement of our present understanding of early arthropod evolution — not least in the highly vexed issue of the 'great appendage problem'[8]. This refers to the controversial identity of a large anterior appendage found in many Cambrian arthropods, and seemingly also in the *Fuxianhuia* specimen described here. Discovering which part of the brain this structure is innervated from will add vital information to this debate. Either way, Ma and colleagues' findings will prompt hasty re-examination of many old specimens, and quite possibly some recasting of recent theories. ■

Graham E. Budd *is in the Department of Earth Sciences, Uppsala University, 75236 Uppsala, Sweden.*
e-mail: graham.budd@pal.uu.se

1. Li, Q.-G. *et al. Science* **335,** 1215–1219 (2012).
2. Martill, D. M. *Nature* **346,** 171–172 (1990).
3. Dong, X.-P. et al. *Nature* **427,** 237–240 (2004).
4. Ma, X.-Y. *et al. Nature* **490,** 258–261 (2012).
5. Sansom, R. S. et al. *Proc. R. Soc. B* **278,** 1150–1157 (2011).
6. Conway Morris, S. *Spec. Papers Palaeont.* **20,** 1–95 (1977).
7. Bergström, J. *et al. GFF* **130,** 189–201 (2008).
8. Budd, G. E. *Palaeontology* **51,** 561–573 (2008).
9. Harzsch, S. *Integr. Comp. Biol.* **46,** 162–194 (2006).


Complete Solution for microRNA

miRNA Profiling
Cancer, disease and miRNome qPCR arrays

miRNA Target Validation
3' UTR target clones with live cell analysis

miRNA Overexpression & Inhibition
Precursor and inhibitor clones and lentiviruses

Choose Your Tools!

Relative Activity Values GLuc/AP
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Gluc-LIN28-3'UTR + + + + + +
miR-125a - + + + + +
NC-Inhibitor - - + - - -
Anti-miR-125a (ng) - - - 20 50 100
Gluc: Gaussia luciferase; AP: Alkaline Phosphatase

GeneCopoeia™
Expressway to Discovery
genecopoeia.com/miRNA
866-360-9531

ANIMAL MIGRATION

Catching the wave

Observations of the migration patterns of Norwegian red deer show that some animals ride waves of greener vegetation as spring spreads across the landscape, whereas others jump ahead in anticipation of this higher-quality food.

JOHN M. FRYXELL & TAL AVGAR

The migration of animals en masse provides one of the great mysteries of life. Birds do it, bugs do it, even fish in the deep blue sea do it. But what is it that compels them to pick up and move with monotonous regularity, giving up one seemingly good home in search of another? Writing in *American Naturalist*, Bischof *et al.*[1] present a field study of red deer (*Cervus elaphus*) (Fig. 1) in Norway that offers a tantalizing hint as to the method behind this phenomenon. Their observations suggest that migration allows deer to take advantage of temporally changing feeding conditions in different parts of the country that are dictated by variation in elevation and latitude.

Such a behavioural strategy would only make sense in a world in which growing conditions are never identical from one place to another. As every gardener knows, this is indeed the case. For example, the onset of the growing season is slower at high latitudes than it is nearer the equator, and plants at high elevation experience spring much later than those at sea level[2]. In a country such as Norway, these geographic realities translate into waves of vegetation 'green-up' that commence at the end of winter and gradually spread northwards and upwards to higher elevations throughout the spring and summer[2,3].

Herbivorous mammals benefit most from feeding on younger, rapidly growing shoots rather than older plants, which are laden with indigestible compounds such as lignin and cellulose[4]. As a result, animals trying to restore weight after an arduous winter, or those shouldering the demands of lactation, could benefit by finding the crest of these green waves as they move across the landscape. Recent studies suggest that some migratory herbivores, such as elk in the Rocky Mountains[3] or gazelles in Mongolia[5], obtain better food than their non-migratory brethren. But what is not clear from previous studies is whether migrant animals simply 'surf' on the peak of a single green wave, continually moving with it over the course of the growing season, or if they 'jump' between peaks (Fig. 2). Because migrations often occur over vast spaces, it is difficult to measure the success of one movement strategy against other options. Bischof and colleagues solved this problem using an elegant form of statistical book-keeping that allowed them to better understand how migrating animals value different spatial locations and to assess how much of an improvement in food intake is achieved by different migration styles.

To tackle this problem, the researchers took advantage of a satellite-derived vegetation index termed NDVI, which is a remote-sensed indicator of plant abundance[2]. Using bi-weekly snapshots of NDVI collected over a period of nine years, they estimated the date of fastest plant growth at each location. Then, for each of 294 study animals, they prepared a matrix that linked plant-growth data with the temporal sequence of the location of the deer. Summed values along the diagonal of each matrix provide a measurement of the quality of food experienced by a given animal over the course of the growing season. The sum of other row and column combinations reflect the food value that would have been obtained with alternative departure dates the deer might have taken. If a particular deer is truly surfing the green wave, then the sum of food qualities along its matrix diagonal should exceed that of any other departure schedule.

Bischof and colleagues differentiated migratory from resident individuals. Their results demonstrate that migratory individuals perform better than resident individuals, who in turn perform better than hypothetical migratory animals would if they had lingered in their winter home ranges rather than migrating. The authors' comparisons of deer from different regions of Norway suggest that the proportion of migratory individuals scales

M. WOIKE/FOTO NATURA/MINDEN PICTURES/CORBIS

Figure 1 | Greener pastures. Some red deer (*Cervus elaphus*) migrate each year to take advantage of new plant growth that emerges with the arrival of spring.

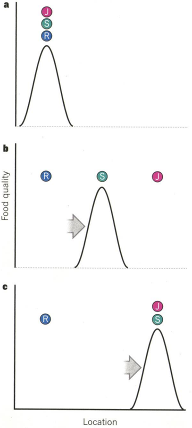


Figure 2 | Surfers and jumpers. Norwegian red deer change their grazing sites in response to the waves of vegetation greening that start at lower, more southerly locations at the end of winter (**a**) and gradually spread northwards and to higher elevations throughout the spring and summer. **b**,**c**, As a green wave travels across the landscape, resident (R) animals lag behind, whereas there are two possible movement strategies for migrating animals: 'surfers' (S) travel with the wave, whereas 'jumpers' (J) move ahead in anticipation of where the wave will be at a later date. Bischof and colleagues' combined analysis[1] of satellite data on vegetation abundance and animal tracking shows that most deer use the 'jumper' strategy.

with the potential gains that are obtainable. Surprisingly, however, they found that most migrant animals do not surf the green wave but rather jump it, arriving in their summer range well before the arrival of optimal feeding conditions. Furthermore, it seems that some animals leap from peak to peak over the course of the season. This suggests that many deer do not follow the optimal surfing strategy, but nonetheless benefit to a lesser degree by periodically jumping into greener pastures. Why they do so remains an open question — perhaps their jumping strategy stems from the need to balance feeding with physiological costs, predation risk or other constraints that might affect their survival during migration[3].

As with all new approaches, there are questionable assumptions in this study. For instance, one would like to have more evidence that the date of fastest plant growth actually translates into maximum food quality. Also, there is some uncertainty surrounding the authors' classifications, because distinguishing between migrants and residents is often troublesome in species that occupy large home ranges, particularly when movements occur as a series of jumps rather than a smooth progression. Perhaps most important is the issue that, although the authors' method allows rigorous assessment of the ecological consequences of different patterns of movement timing along a single trajectory, it cannot tell us the relative value of a particular spatial trajectory against the infinite number of other conceivable patterns. This challenge invokes the well-known combinatorial problem referred to as the travelling salesman dilemma[6]. Although such complex spatial problems currently cannot be solved exactly, new statistical modelling techniques based on likelihood assessments could help to provide robust estimates[7].

These complications aside, Bischof and colleagues' study offers a fresh take on movement ecology, a burgeoning field that blends developments in movement modelling with technological advances made available by satellite-based information systems[8]. The logic used by the authors should, in principle, be applicable to a broad range of animal movement patterns, ranging from nomadism to territoriality. Perhaps the day is not far off when the complex pattern of steps taken by an individual over its entire life can be predicted on the basis of the continually shifting mosaic of resources and costs with which it is confronted. ■

John M. Fryxell *and* **Tal Avgar** *are in the Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada.*
e-mails: jfryxell@uoguelph.ca; tavgar@uoguelph.ca

1. Bischof, R. *et al. Am. Nat.* **180,** 407–424 (2012).
2. Pettorelli, N. *et al. Trends Ecol. Evol.* **20,** 503–510 (2005).
3. Hebblewhite, M., Merrill, E. & McDermid, G. *Ecol. Monogr.* **78,** 141–166 (2008).
4. Langvatn, R. & Hanley, T. A. *Oecologia* **95,** 164–170 (1993).
5. Mueller, T. *et al. J. Appl. Ecol.* **45,** 649–658 (2008).
6. Anderson, D. J. *Theor. Popul. Biol.* **24,** 145–159 (1983).
7. Patterson, T. A. *et al. Trends Ecol. Evol.* **23,** 87–94 (2008).
8. Nathan, R. *et al. Proc. Natl Acad. Sci. USA* **105,** 19052–19059 (2008).

EARTH SCIENCE

When an oceanic tectonic plate cracks

Analyses of two recent earthquakes of great magnitude show how complex the breaking of the oceanic lithosphere can be, how it is linked to earlier great events and how it triggers seismicity worldwide. SEE LETTERS P.240, P.245 & P.250

JEAN-YVES ROYER

On 11 April 2012, two great earthquakes of magnitude (M_w) 8.6 and 8.2 struck the northeastern Indian Ocean, a few hundred kilometres off the Sunda Trench that lies just west of Indonesia (Fig. 1). The two earthquakes were close in time and space (2 hours and 185 km apart), and occurred not far from the epicentres of the devastating megathrust earthquakes that hit Sumatra's Aceh region in 2004 (M_w 9.2) and Nias Island in 2005 (M_w 8.7). Fortunately, the 2012 intra-oceanic events were not damaging to human activities, but they will be marked in the annals of seismology for several reasons: first, because of their unusually high magnitudes given that they occurred in the interior of an oceanic tectonic plate, rather than at a plate boundary; second, because the 'strike-slip' faulting mechanism of the events is unexpected for earthquakes of such magnitude; and finally, because of the complexity of the rupture of the intervening plate. These exceptional properties put the M_w-8.6 mainshock into the top-ten list of the world's largest earthquakes since 1900, and promptly aroused the curiosity of seismologists[1,2] and those working on plate tectonics, because the events took place in one of the largest and most complex deforming zones in the oceanic lithosphere (Earth's crust and uppermost mantle). Three papers in this issue[3–5] investigate these unprecedented events.

Yue *et al.*[4] (page 245) unravel the extraordinary complexity of the underlying mechanisms of the 2012 events, which display perpendicular (orthogonal) and discontinuous (*en échelon*) fault ruptures. By analysing short-period and long-period seismic waves recorded by stations in Europe and Japan, the authors identified a sequence of shocks along four fault planes, and estimated the rupture length and amount of slip along the faults. The deduced radiation pattern of the seismic waves outlines the interplay between multiple orthogonal and *en échelon* faults, whose spatial distribution is consistent with the epicentres of the first-week aftershocks. The orientations of the faults reflect the seafloor's tectonic fabric[2,4], with ruptures occurring alternately along faults that extend bilaterally west-northwest and east-southeast (parallel to the abyssal hills) or along faults extending north-northeast and south-southwest (parallel to the fracture zone's orientation), the latter being perpendicular to the former. The high-level seismicity in the Wharton Basin (Fig. 1), as well as seafloor mapping, had already provided direct evidence that the fossil fracture zones — seismically inactive, long linear faults generated by seafloor spreading — in this region are continually reactivated[6–9]. But the question remains as to why so much stress can accumulate and provoke such large-magnitude earthquakes, and furthermore in an area that could have been activated by the nearby Aceh megathrust earthquake.

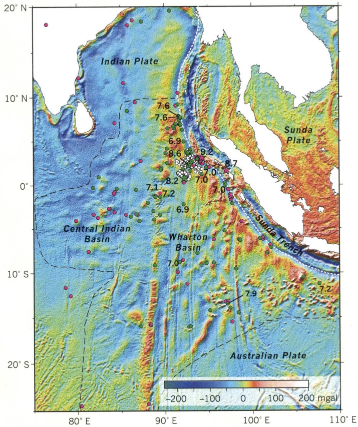


Figure 1 | Seismo-tectonic chart of the India–Australia plate. Earthquakes in this region have either strike-slip faulting mechanisms (green dots and stars) with faults oriented parallel to fracture zones, or thrust faulting mechanisms (red dots and stars) with faults generally perpendicular to the fracture zones[6,13]. Two large earthquakes (green stars) that occurred on 11 April 2012 have now been investigated[3–5]; aftershocks of those earthquakes are shown in white circles[14]. The magnitudes (specifically the 'moment' magnitudes, M_w) of earthquakes equal to or greater than 6.9 are indicated. Seismic events related to the subduction zone (east of the Sunda Trench, which is depicted as a white, broken line) are not shown, except for the megathrust earthquakes (red stars) at Aceh in 2004 (M_w 9.2) and at Nias in 2005 (M_w 8.7). Events with unknown fault mechanisms within the India–Australia plate boundary are also not shown. Black dashed lines define the limits of the active boundary between the Indian and Australian plates[11]. Gravity anomalies (measured in mgal) reflect the ocean basin's topography beneath sediments, and are indicated by colours on the map: troughs are associated with negative anomalies, and ridges with positive ones. There is a close relationship between the distribution of earthquakes, particularly those that occurred on 11 April 2012, and the extinct seafloor fabric of the eastern Indian Ocean, such as the long north–south oriented fracture zones (lineated troughs) in the Wharton Basin.

This is exactly what Delescluse *et al.*[3] (page 240) attempt to answer, by investigating the causal relationship between the high-level seismicity along the Sunda Trench (Fig. 1) — in particular the Aceh and Nias megathrust events — and the off-trench intra-oceanic seismicity. From Delescluse and colleagues' Figure 2 (page 241), there is no doubt that the Aceh and Nias events have shaken the Indo–Australian lithosphere which is plunging (subducting) beneath the Sunda plate, and that they triggered a long-lasting, large increase in seismic activity off-trench. By looking at how stresses are released in the vicinity of large ruptures (Coulomb stress changes), the authors show that co-seismic slips of the Aceh and Nias megathrusts can promote left-lateral strike-slip earthquakes with a direction parallel to the orientation of the Wharton Basin fracture zones, similar to the mechanisms observed in the April 2012 events. However, such an approach assumes an elastic rheology, for which stress changes decrease immediately after the earthquake, as a new interseismic stress-loading cycle starts along ruptured faults. Delescluse *et al.* explain the seven-year lag between the Aceh and April 2012 events by a buffering effect of the viscoelastic properties of the asthenosphere, the region of the mantle that underlies the lithosphere. The authors' model predicts that the maximum post-seismic stresses generated by the Aceh megathrust could have reached the area of the April 2012 earthquakes seven to ten years later.

In turn, the two April 2012 events may have triggered seismicity worldwide. Pollitz *et al.*[5] (page 250) show that the number of remote earthquakes with magnitudes larger than 5.5 and up to 7 increased worldwide nearly fivefold in the six days following the April 2012 events. The authors attribute this effect to the powerful radiation pattern of surface seismic waves generated by the strike-slip mechanisms, but cannot exclude the possibility that the 2012 earthquakes struck during a particularly quiet seismic period, and so triggered an unusually large number of earthquakes that were already on the brink of occurring.

Interestingly, the papers by Delescluse *et al.* and by Pollitz *et al.* reveal contrasting short-term and long-term 'domino effects' of major seismic events: the April 2012 events may have triggered quasi-immediate aftershocks far away, whereas the 2004 and 2005 events possibly triggered aftershocks nearby several years later. The two studies emphasize how the rupture mechanism and the tectonic setting of earthquakes govern their ability to trigger further seismic activity through seismic-wave excitation and short-term or long-term stress relaxation

The occurrence off-trench of such major events is perhaps better understood in the context of the high lithospheric stresses caused by differential motion between the Indian plate, which is slowed down in its northward drift by collision with the Eurasian plate, and the Australian plate, which is rapidly subducting underneath Indonesia along the Sunda Trench. As a result, the two plates move towards each other at a rate of 15 millimetres per year in an east-southeast direction[10], and have been doing so for at least 10 million years[11]. The April 2012 events have the largest magnitudes ever recorded in the diffuse (as opposed to narrow) India–Australia plate boundary, where earthquakes of magnitude 7 and greater are not uncommon (Fig. 1).

Do these major intraplate earthquakes indicate the birth of a new, discrete plate boundary? Probably not, because this unusually high-level seismicity occurs over large areas (diffuse boundaries) bounded by otherwise stable and 'quiet' oceanic plate interiors (rigid plates), with all these pieces constituting the composite, deforming Indo–Australian lithosphere[11]. The April 2012 events illustrate how oceanic lithosphere can fail in complex ways[2,4,9] as a result of the fabric structure it has inherited from seafloor spreading[12], and they shed new light on its mechanical behaviour and strength, and on earthquake physics. ■

Jean-Yves Royer *is at the Laboratoire Domaines Océaniques, Institute for Marine Studies, CNRS and University of Brest, 29280 Plouzané, France.*
e-mail: jean-yves.royer@univ-brest.fr

1. McGuire, J. J. & Beroza, G. C. *Science* **336,** 1118–1119 (2012).
2. Meng, L. *et al. Science* **337,** 724–726 (2012).
3. Delescluse, M. *et al. Nature* **490,** 240–244 (2012).
4. Yue, H., Lay, T. & Koper, K. D. *Nature* **490,** 245–249 (2012).
5. Pollitz, F. F. et al. *Nature* **490,** 250–253 (2012).
6. Petroy, D. E. & Wiens, D. A. *J. Geophys. Res.* **94,** 12301–12319 (1989).
7. Deplus, C. *et al. Geology* **26,** 131–134 (1998).
8. Robinson, D. P. et al. *Science* **292,** 1145–1148 (2001).
9. Abercrombie, R. E., Antolik, M. & Ekström, G. *J. Geophys. Res.* **108,** 16–31 (2003).
10. Gordon, R. G. *et al. Geology* **36,** 827–830 (2008).
11. Royer, J.-Y. & Gordon, R. G. *Science* **277,** 1268–1274 (1997).
12. Delescluse, M. et al. *Geophys. Res. Lett.* **35,** L16312 (2008).
13. www.globalcmt.org
14. http://earthquake.usgs.gov

BRAIN DEVELOPMENT

The neuron family tree remodelled

The discovery of different classes of neuronal progenitor cell, destined to give rise to neurons in specific layers of the cerebral cortex, could presage the revision of a 50-year-old model of brain development.

OSCAR MARÍN

The neocortex, the mammalian brain's most recent evolutionary acquisition, controls the aspects of behaviour that make us human, from the fine-scale finger movements of a virtuoso pianist to the complex syntactic processing required for language. Its highly regular structure contains a complex matrix of excitatory and inhibitory neurons organized into distinct layers and columns. Neurons in any one cortical layer share general patterns of connectivity, whereas neurons in the same column are typically interconnected across layers and function as the basic unit of cortical operations[1]. Over the past 100 years, analysis of the developing brain has provided fundamental insights into the functional organization of the neocortex[2]. However, Franco *et al.*[3], writing in *Science*, provide evidence that puts a radically new perspective on the link between brain development and neuronal function.

A fascinating aspect of cortical development is that excitatory neurons (also known as pyramidal cells) are 'born' in sequential order, with those located in deep layers of the neocortex being generated first and subsequently generated neurons being positioned in progressively more superficial layers. It is usually assumed that pyramidal cells in all layers of the neocortex originate from a single type of progenitor cell, and that a progenitor's ability to generate distinct classes of neuron decreases with time. That is, progenitors during early embryonic development can give rise to any class of pyramidal cell (they are 'multipotent') but are progressively restricted to producing superficial neurons later on[4–6]. New neurons, originating from transient amplifying cells that become detached from the ventricular zone (the inner layer of the cortex, where progenitors reside), migrate to their final position by following the radial fibres of progenitor cells (Fig. 1a). These fibres serve as a scaffold that links the ventricular zone and the cortical surface, just like spokes connecting the centre of a wheel to its outer edge.

According to this classical view, the 'birth date' of a neuron largely determines its fate, and sibling neurons are vertically aligned through the radial axis of the neocortex. This model has been highly influential, because it relates the concept of ontogenetic columns (vertical arrangements of sibling neurons born from a common progenitor) with the notion of functional columns — the long-sought fundamental unit of cortical computation. Indeed, recent studies suggest that sibling neurons are more likely to respond to the same sensory stimuli than a random subset of neighbouring neurons[7,8].

Franco *et al.* report that the gene-regulatory protein Cux2 is expressed by a small subset of progenitor cells in the developing cortex of mice, and that this subset increases over time. These findings are unexpected, because Cux2 expression was previously thought to be restricted to pyramidal cells in the superficial layers of the cortex. The authors then hypothesized that the Cux2-expressing progenitor cells might be fated to generate superficial pyramidal cells. To test this possibility, they generated and studied a mouse strain in which a fluorescent protein was produced exclusively in cells

50 Years Ago

It is now fifty years since, with the creation of the former Medical Research Committee, Great Britain took the lead in setting up central organizations for research ... The Medical Research Council conceives its function as: to watch over the whole fields of medical and related biological research so as to foresee, to the best of its ability, the needs and opportunities to give support to any promising research in these fields irrespective of the agent concerned; to work in partnership with the universities and professions on one hand and the various Government departments on the other, so that new knowledge may be made available as the need arises.
From *Nature* 13 October 1962

100 Years Ago

Coming out from Queenstown on September 10 on her way toward Boston, the ss. *Arabic* was accompanied for some hours by a large flock of gulls. For the most part these birds were visibly working, flapping their wings, but occasionally a few would cease flapping and merely sail along for considerable distances ... An explanation of the ability of the birds to sail, under the conditions described above, is, I believe, found in the upward course of the wind which has struck the weather side of the ship and must rise in order to pass over it ... As the trail of smoke marks the direction of the wind with respect to the moving ship, the bird must, in order to sail with the same velocity and direction as the ship, have a motion relative to the air equal and opposite to the motion of the smoke relative to the ship. Accordingly, the bird's axis is kept parallel to, and opposite to, the course of the smoke as indicated by its train from the funnel.
From *Nature* 10 October 1912

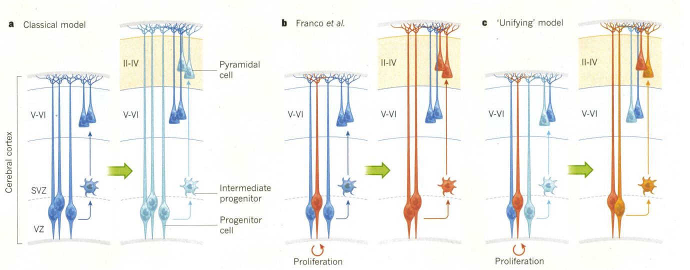


Figure 1 | Adding layers of complexity to the brain. Several models have been proposed to explain the formation of neural layers (II to VI) during the development of the cerebral cortex. **a**, The classical model suggests that there is a homogeneous population of multipotent progenitor cells in the ventricular zone (VZ) of the cortex that produce intermediate, transient, dividing progenitors located in the subventricular zone (SVZ), which in turn give rise to different classes of pyramidal cell (a type of neuron) in sequential order. According to this model, progenitors can generate any class of pyramidal cell during early development, but are progressively restricted to producing superficial neurons at later stages. **b**, Franco and colleagues' findings[3] indicate that there are two classes of progenitor cell, each with a defined fate: one class gives rise to deep pyramidal cells (layers V and VI), whereas the other generates superficial pyramidal cells (layers II to IV). During early development, only a few progenitor cells exist that are destined to form superficial layers; but these progenitors proliferate before they begin producing neurons. **c**, In a unifying hypothesis that attempts to reconcile different sets of observations, several progenitor lineages exist. Progenitor cells are fate-restricted to generate deep- or superficial-layer neurons, but, within each lineage, intrinsic or extrinsic factors modulate the production of distinct classes of pyramidal cell (dark and light colours), which are vertically aligned in progressively more superficial positions.

that either expressed Cux2 (Cux2$^+$ cells) or were derived from Cux2-expressing cells. Using this lineage-tracing technique, the researchers found that most neurons generated from Cux2$^+$ progenitor cells settle in the superficial layers of the cortex (layers II to IV), whereas those derived from Cux2$^-$ progenitors end up in the deep cortical layers (layers V and VI).

To gain further support for their hypothesis, the authors created a second mouse strain in which the fluorescent labelling of Cux2$^+$ cells could be triggered by the injection of a specific compound (tamoxifen) into pregnant mice at different times during embryonic development. This approach allowed Franco *et al.* to observe the unequivocal labelling of Cux2$^+$ progenitor cells at very early stages of brain development, and to confirm that Cux2$^+$ progenitor cells give rise almost exclusively to pyramidal cells located in the superficial layers of the cortex.

Taken together, the results show that at least two classes of progenitor cell exist in the neocortex: Cux2$^-$ progenitors generate deep-layer pyramidal cells, whereas Cux2$^+$ progenitors primarily produce pyramidal cells for the superficial layers. Franco and colleagues' results therefore indicate that cortical progenitor cells are programmed to produce specific classes of excitatory neuron (Fig. 1b), a finding that is at odds with the classical view that multipotent progenitor cells exist in the neocortex. According to the authors' findings, distinct classes of pyramidal cell are produced at different times during development because fate-restricted progenitor cells begin to produce neurons at different stages, and not because the potential of the progenitor cells becomes more limited with time. This model is consistent with recent research suggesting that progenitor cells in the cortex are much more diverse than previously thought[9].

However, the model is in conflict with several previous studies. For example, single progenitor cells isolated from the embryonic cortex, or embryonic stem cells that are induced to become cortical progenitors *in vitro*, are multipotent and generate excitatory neurons in the normal *in vivo* order[10,11]. One possible explanation for the conflict is that these older observations reflect the potential of cortical progenitor cells before they become fate-locked at a certain stage of development. In addition, *in vivo* studies using other lineage-tracing methods have shown that pyramidal cells derived from a single progenitor cell are typically arranged in vertical clusters, although mostly within adjacent layers[6,7]. These observations are consistent with a hybrid, 'unifying' model of cortical neurogenesis (also suggested by Franco and colleagues), in which intrinsic or extrinsic factors affect the potential of fate-restricted progenitors to generate specific classes of neuron during brain development (Fig. 1c).

Our current view of the mechanisms that control the assembly of cortical circuits may also need adjustment. For example, if most pyramidal cells from deep and superficial layers are derived from different progenitors, it is unlikely that many pyramidal cells are connected across layers according to lineage relationship[7,8]. Inhibitory interneurons, which modulate the function of pyramidal cells and are derived from progenitor cells outside the cortex[12], have also been suggested to be arranged in relation to the columnar organization of pyramidal cells[13]. However, if Franco and colleagues' results reflect a general property of cortical development, interneuron progenitors might instead be fate-restricted to preferentially match the laminar pattern of the neocortex. Whatever the answers to these questions, I believe that this study marks a major shift in our understanding of cortical development. ■

Oscar Marín *is at the Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas and Universidad Miguel Hernández, Sant Joan d'Alacant 03550, Spain.*
e-mail: o.marin@umh.es

1. Mountcastle, V. B. *Brain* **120,** 701–722 (1997).
2. Rakic, P. *Cereb. Cortex* **16,** i3–i17 (2006).
3. Franco, S. J. *et al. Science* **337,** 746–749 (2012).
4. Rakic, P. *Science* **241,** 170–176 (1988).
5. McConnell, S. K. *Trends Neurosci.* **12,** 342–349 (1989).
6. Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. *Nature* **409,** 714–720 (2001).
7. Li, Y. *et al. Nature* **486,** 118–121 (2012).
8. Ohtsuki, G. *et al. Neuron* **75,** 65–72 (2012).
9. Fietz, S. A. & Huttner, W. B. *Curr. Opin. Neurobiol.* **21,** 23–35 (2011).
10. Shen, Q. *et al. Nature Neurosci.* **9,** 743–751 (2006).
11. Gaspard, N. *et al. Nature* **455,** 351–357 (2008).
12. Anderson, S. A. *et al. Science* **278,** 474–476 (1997).
13. Brown, K. N. *et al. Science* **334,** 480–486 (2011).

doi:10.1038/nature11556

A call for transparent reporting to optimize the predictive value of preclinical research

Story C. Landis[1], Susan G. Amara[2], Khusru Asadullah[3], Chris P. Austin[4], Robi Blumenstein[5], Eileen W. Bradley[6], Ronald G. Crystal[7], Robert B. Darnell[8], Robert J. Ferrante[9], Howard Fillit[10], Robert Finkelstein[1], Marc Fisher[11], Howard E. Gendelman[12], Robert M. Golub[13], John L. Goudreau[14], Robert A. Gross[15], Amelie K. Gubitz[1], Sharon E. Hesterlee[16], David W. Howells[17], John Huguenard[18], Katrina Kelner[19], Walter Koroshetz[1], Dimitri Krainc[20], Stanley E. Lazic[21], Michael S. Levine[22], Malcolm R. Macleod[23], John M. McCall[24], Richard T. Moxley III[25], Kalyani Narasimhan[26], Linda J. Noble[27], Steve Perrin[28], John D. Porter[1], Oswald Steward[29], Ellis Unger[30], Ursula Utz[1] & Shai D. Silberberg[1]

The US National Institute of Neurological Disorders and Stroke convened major stakeholders in June 2012 to discuss how to improve the methodological reporting of animal studies in grant applications and publications. The main workshop recommendation is that at a minimum studies should report on sample-size estimation, whether and how animals were randomized, whether investigators were blind to the treatment, and the handling of data. We recognize that achieving a meaningful improvement in the quality of reporting will require a concerted effort by investigators, reviewers, funding agencies and journal editors. Requiring better reporting of animal studies will raise awareness of the importance of rigorous study design to accelerate scientific progress.

Dissemination of knowledge is the engine that drives scientific progress. Because advances hinge primarily on previous observations, it is essential that studies are reported in sufficient detail to allow the scientific community, research funding agencies and disease advocacy organizations to evaluate the reliability of previous findings. Numerous publications have called attention to the lack of transparency in reporting, yet studies in the life sciences in general, and in animals in particular, still often lack adequate reporting on the design, conduct and analysis of the experiments. To develop a plan for addressing this critical issue, the US National Institute of Neurological Disorders and Stroke (NINDS) convened academic researchers and educators, reviewers, journal editors and representatives from funding agencies, disease advocacy communities and the pharmaceutical industry to discuss the causes of deficient reporting and how they can be addressed. The specific goal of the meeting was to develop recommendations for improving how the results of animal research are reported in manuscripts and grant applications. There was broad agreement that: (1) poor reporting, often associated with poor experimental design, is a significant issue across the life sciences; (2) a core set of research parameters exist that should be addressed when reporting the results of animal experiments; and (3) a concerted effort by all stakeholders, including funding agencies and journals, will be necessary to disseminate and implement best reporting practices throughout the research community. Here we describe the impetus for the meeting and the specific recommendations that were generated.

Widespread deficiencies in methods reporting

In the life sciences, animals are used to elucidate normal biology, to improve understanding of disease pathogenesis, and to develop therapeutic interventions. Animal models are valuable, provided that experiments employing them are carefully designed, interpreted and reported. Several recent articles, commentaries and editorials highlight that inadequate experimental reporting can result in such studies being un-interpretable and difficult to reproduce[1–8]. For instance, replication of spinal cord injury studies through an NINDS-funded program determined that many studies could not be replicated because of incomplete or inaccurate description of experimental design, especially how randomization of animals to the various test groups, group formulation and delineation of animal attrition and exclusion were addressed[7]. A review of 100 articles published in *Cancer Research* in 2010 revealed that only 28% of papers reported that animals were randomly allocated to treatment groups, just 2% of papers reported that observers were blinded to treatment, and none stated the methods used to determine the number of animals per group, a determination required to avoid false outcomes[2]. In addition, analysis of several

---

[1]National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland 20892, USA. [2]Department of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213, USA. [3]Bayer HealthCare, 13342 Berlin, Germany. [4]National Center for Advancing Translational Sciences, NIH, Rockville, Maryland 20854, USA. [5]CHDI Management/CHDI Foundation, New York, New York 10001, USA. [6]Center for Review, NIH, Bethesda, Maryland 20892, USA. [7]Department of Genetic Medicine, Weill Cornell Medical College, New York, New York 10021, USA. [8]Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10065, USA. [9]Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [10]Alzheimer's Drug Discovery Foundation, New York, New York 10019, USA. [11]Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts 01545, USA. [12]Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, Nebraska 68198, USA. [13]JAMA, Chicago, Illinois 60654, USA. [14]Department of Neurology, Michigan State University, East Lansing, Michigan 48824, USA. [15]Department of Neurology, University of Rochester Medical Center, Rochester, New York 14642, USA. [16]Parent Project Muscular Dystrophy, Hackensack, New Jersey 07601, USA. [17]The Florey Institute of Neuroscience and Mental Health, University of Melbourne, Heidelberg 3081, Australia. [18]Neurology and Neurological Sciences and Cellular and Molecular Physiology, Stanford University, Stanford, California 94305, USA. [19]Science Translational Medicine, AAAS, Washington DC 22201, USA. [20]Department of Neurology, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [21]F. Hoffmann-La Roche, 4070 Basel, Switzerland. [22]Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, California 90095, USA. [23]Department of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK. [24]PharMac LLC, Boca Grande, Florida 33921, USA. [25]University of Rochester Medical Center, School of Medicine and Dentistry, Rochester, New York 14642, USA. [26]Nature Neuroscience, New York, New York 10013, USA. [27]Department of Neurological Surgery, University of California San Francisco, San Francisco, California 94143, USA. [28]ALS Therapy Development Institute, Cambridge, Massachusetts 02139, USA. [29]Reeve-Irvine Research Center, University of California Irvine, Irvine, California 92697, USA. [30]Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland 20993, USA.

hundred studies conducted in animal models of stroke, Parkinson's disease and multiple sclerosis also revealed deficiencies in reporting key methodological parameters that can introduce bias[6]. Similarly, a review of 76 high-impact (cited more than 500 times) animal studies showed that the publications lacked descriptions of crucial methodological information that would allow informed judgment about the findings[9]. These deficiencies in the reporting of animal study design, which are clearly widespread, raise the concern that the reviewers of these studies could not adequately identify potential limitations in the experimental design and/or data analysis, limiting the benefit of the findings.

Some poorly reported studies may in fact be well-designed and well-conducted, but analysis suggests that inadequate reporting correlates with overstated findings[10–14]. Problems related to inadequate study design surfaced early in the stroke research community, as investigators tried to understand why multiple clinical trials based on positive results in animal studies ultimately failed. Part of the problem is, of course, that no animal model can fully reproduce all the features of human stroke. It also became clear, however, that many of the difficulties stemmed from a lack of methodological rigor in the preclinical studies that were not adequately reported[15]. For instance, a systematic review and meta-analysis of studies testing the efficacy of the free-radical scavenger NXY-059 in models of ischaemic stroke revealed that publications that included information on randomization, concealment of group allocation, or blinded assessment of outcomes reported significantly smaller effect sizes of NXY-059 in comparison to studies lacking this information[10]. In certain cases, a series of poorly designed studies, obscured by deficient reporting, may, in aggregate, serve erroneously as the scientific rationale for large, expensive and ultimately unsuccessful clinical trials. Such trials may unnecessarily expose patients to potentially harmful agents, prevent these patients from participating in other trials of possibly effective agents, and drain valuable resources and energy that might otherwise be more productively spent.

A core set of reporting standards

The large fraction of poorly reported animal studies and the empirical evidence of associated bias[6,10–14,16–20], defined broadly as the introduction of an unintentional difference between comparison groups, led various disease communities to adopt general[21–23] and animal-model-specific[6,24–26] reporting guidelines. However, for guidelines to be effective and broadly accepted by all stakeholders, they should be universal and focus on widely accepted core issues that are important for study evaluation. Therefore, based on available data, we recommend that, at minimum, authors of grant applications and scientific publications should report on randomization, blinding, sample-size estimation and the handling of all data (see below and Box 1).

Randomization and blinding

Choices made by investigators during the design, conduct and interpretation of experiments can introduce bias, resulting in false-positive results. Many have emphasized the importance of randomization and blinding as a means to reduce bias[6,21–23,27], yet inadequate reporting of these aspects of study design remains widespread in preclinical research. It is important to report whether the allocation, treatment and handling of animals were the same across study groups. The selection and source of control animals needs to be reported as well, including whether they are true littermates of the test groups. Best practices should also include reporting on the methods of animal randomization to the various experimental groups, as well as on random (or appropriately blocked) sample processing and collection of data. Attention to these details will avoid mistaking batch effects for treatment effects (for example, dividing samples from a large study into multiple lots, which are then processed separately). Investigators should also report on whether the individuals caring for the animals and conducting the experiments were blinded to the allocation sequence, blinded to group allocation and, whenever possible, whether the persons assessing, measuring or quantifying the experimental outcomes were blinded to the intervention.

BOX 1

A core set of reporting standards for rigorous study design

Randomization
- Animals should be assigned randomly to the various experimental groups, and the method of randomization reported.
- Data should be collected and processed randomly or appropriately blocked.

Blinding
- Allocation concealment: the investigator should be unaware of the group to which the next animal taken from a cage will be allocated.
- Blinded conduct of the experiment: animal caretakers and investigators conducting the experiments should be blinded to the allocation sequence.
- Blinded assessment of outcome: investigators assessing, measuring or quantifying experimental outcomes should be blinded to the intervention.

Sample-size estimation
- An appropriate sample size should be computed when the study is being designed and the statistical method of computation reported.
- Statistical methods that take into account multiple evaluations of the data should be used when an interim evaluation is carried out.

Data handling
- Rules for stopping data collection should be defined in advance.
- Criteria for inclusion and exclusion of data should be established prospectively.
- How outliers will be defined and handled should be decided when the experiment is being designed, and any data removed before analysis should be reported.
- The primary end point should be prospectively selected. If multiple end points are to be assessed, then appropriate statistical corrections should be applied.
- Investigators should report on data missing because of attrition or exclusion.
- Pseudo replicate issues need to be considered during study design and analysis.
- Investigators should report how often a particular experiment was performed and whether results were substantiated by repetition under a range of conditions.

Sample-size estimation

Minimizing the use of animals in research is not only a requirement of funding agencies around the world but also an ethical obligation. It is unethical, however, to perform underpowered experiments with insufficient numbers of animals that have little prospect of detecting meaningful differences between groups. In addition, with smaller studies, the positive predictive value is lower, and false-positive results can ensue, leading to the needless use of animals in subsequent studies that build upon the incorrect results[28]. Studies with an inadequate sample size may also provide false-negative results, where potentially important findings go undetected. For these reasons it is crucial to report how many animals were used per group and what statistical methods were used to determine this number.

Data handling

Common practices related to data handling that can also lead to false positives include interim data analysis[29], the *ad hoc* exclusion of data[30], retrospective primary end point selection[31], pseudo replication[32] and small effect sizes[33].

Interim data analysis

It is not uncommon for investigators to collect some data and perform an interim data analysis. If the results are statistically significant in favour of the working hypothesis, the study is terminated and a paper is written. If the results look 'promising' but are not statistically significant, additional data are collected. This has been referred to as 'sampling to a foregone conclusion' and can lead to a high rate of false-positive findings[29,30]. Therefore, sample size and rules for stopping data collection should be defined in advance and properly reported. Unplanned interim analyses, which can inflate false-positive outcomes and require unblinding of the allocation code, should be avoided. If there are interim analyses, however, these should be reported in the publication.

Ad hoc exclusion of data

Animal studies are often complex and outliers are not unusual. Decisions to include or exclude specific animals on the basis of outcomes (for example, state of health, dissimilarity to other data) have the potential to influence the study results. Thus, rules for inclusion and exclusion of data should be defined prospectively and reported. It is also important to report whether all animals that were entered into the experiment actually completed it, or whether they were removed, and if so, for what reason. Differential attrition between groups can introduce bias. For example, a treatment may appear effective if it kills off the weakest or most severely affected animals whose fates are then not reported. In addition, it is important to report whether any data were removed before analysis and the reasons for this data exclusion.

Retrospective primary end-point selection

It is well known that assessment of multiple end points, and/or assessment of a single end point at multiple time points, inflates the type-I error (false-positive results)[31]. Yet it is not uncommon for investigators to select a primary end point only after data analyses. False-positive conclusions arising from such practices can be avoided by specifying a primary end point before the study is undertaken, the time(s) at which the end point will be assessed, and the method(s) of analysis. Significant findings for secondary end points can and should be reported, but should be delineated as exploratory in nature. If multiple end points are to be assessed, then appropriate statistical corrections should be applied to control type-I error, such as Bonferroni corrections[31,34].

Pseudo replicates

When considering sample-size determination and experimental design, pseudo-replication issues need to be considered[32]. There is a clear, but often misunderstood or misrepresented, distinction between technical and biologic replicates. For example, in analysing effects of pollutants on reproductive health, multiple sampling from a litter, regardless of how many littermates are quantified, provides data from only a single biologic replicate. When biologic variation in response to some intervention is the variable of interest, as in many animal experiments, analysis of samples from multiple litters is essential. The unit of assessment is the smallest unit (animal, cage, litter) to which the intervention in question can be independently administered[35].

Small effect sizes

A statistically significant result does not provide information on the magnitude of the effect and thus does not necessarily mean that the effect is robust, which could account for the poor reproducibility of certain studies[36]. Therefore, reporting whether results were substantiated by repetition, preferably under a range of conditions that demonstrate the robustness of the effect is encouraged. Also, reporting how often the particular experiment was performed as a means to control for a general tendency to publish only the best results would strengthen the validity of experimental results. To this end, carefully designed and powered animal studies should be budgeted for in the grant applications and funding agencies should consider supporting repetition studies where appropriate.

An important note about exploratory experiments

For the most part, these best practices do not apply to early-stage observational experiments searching for possible differences among experimental groups. Such exploratory testing is frequently conducted using a small sample size, does not have a primary outcome and is often unblinded. However, because such experiments are likely to be subject to many of the limitations described above, they should be viewed as hypothesis-generating experiments and interpreted as such. Potential discoveries arising from the exploratory phase of the research should be supported by follow-up, hypothesis-testing experiments that take into consideration and adequately report on the core standards detailed above (Box 1).

The path to implementation

Improving the transparency and quality of reporting cannot be achieved by a single party, but will require cooperation among all stakeholders, including investigators, reviewers, funding agencies and journals. Calling upon investigators to provide key information about the design, execution and analysis of animal experiments described in grant applications and manuscripts and encouraging reviewers to consider these issues in their evaluations should, over time, increase both the quality and predictive value of preclinical research. Potential strategies for achieving this goal can be adopted from the clinical trials community, which also contended with poor reporting and associated bias. Evidence that clinical trials can yield biased results if they lack methodological rigor[37–42] led to the development and implementation of the CONSORT guidelines for randomized clinical trials (among other guidelines), now adopted by many clinical journals and funding organizations. These guidelines require that authors report whether and how their studies were carried out blind and randomized, how sample size was determined, whether data are missing owing to attrition or exclusion, and supply information about other important experimental parameters[43–45]. Importantly, the guidelines have improved the transparency of clinical study reporting in journals that have adopted them[46–49]. Additional evidence for the power of such guidelines can be deduced from the observation that, although few animal studies report on randomization, blinding or sample-size determination, most describe compliance with animal regulations, which is required by journals[6,9,10,50,51].

As a first step, we recommend that funding organizations and journals provide reviewers with clear guidance about core features of animal study design (listed in Box 1). The goal is not to be prescriptive or proscriptive, but rather to delineate the minimum set of standards that should routinely be considered in evaluating the appropriateness of a study. Such guidance would make the task easier for reviewers of manuscripts and grant applications who volunteer their time and are often overextended. In addition, investigators and reviewers should be encouraged to consult published generic and model-specific guidelines for designing *in vivo* animal experiments[6,21–27,52,53]. To assist reviewers, editors and funding organizations in making sure that applications and manuscripts contain sufficient information on the core reporting recommendation (Box 1), authors could be asked to append relevant information on a standardized form that accompanies the submission. This form could be as simple as a checkbox indicating the page on which the key reporting standard is addressed. Such a form is currently used by clinical research journals.

In addition to the measures proposed above, better dissemination of knowledge will be greatly facilitated by addressing publication bias, the phenomenon that few studies showing negative outcomes are published[54–63]. Such deficiency in reporting contributes to needless repetition of similar studies by investigators unaware of earlier efforts[60,61]. There is a widely accepted belief that the scientific community, promotions committees, funding agencies and journals favour positive outcomes, an impression that can lead to bias[64]. Possible solutions include incentivizing investigators to publish negative outcomes, supporting studies of independent replication, encouraging journals to publish a greater number of studies reporting negative outcomes, creating a database for negative outcomes (analogous to http://ClinicalTrials.gov/), and linking the raw data to publications.

Change will not occur overnight. The importance of training scientists to properly design and adequately report animal studies cannot be overstated. Training and education focused on key features of experimental design should be an ongoing process for both the novice and veteran involved in biomedical research. Attention to better study design reporting should be communicated at major meetings, brought to the attention of reviewers, editors and funders, required by the publishers of peer-review journals, and included in the training program of graduate and postdoctoral students. Furthermore, good mentorship is crucial for developing such skills and should be encouraged and rewarded. Rigorous experimental design and adequate reporting needs to be emphasized across the board and monitored in training grants awarded by the US National Institute of Health (NIH) and other funding agencies. Professional societies can also have an important role by highlighting this issue in their respective communities.

An important gatekeeper of quality remains the peer review of grant applications and journal manuscripts. We therefore call upon funding agencies and publishing groups to take actions to reinforce the importance of methodological rigor and reporting. NINDS has begun taking steps to promote best practices for preclinical therapy development studies. In 2011, a Notice was published in the NIH Guide encouraging the scientific community to address the issues described above in their grant applications, in describing both the project being proposed and the supporting data upon which it is based (http://grants.nih.gov/grants/guide/notice-files/NOT-NS-11-023.html). Points that should be considered in a well-designed study are listed on the NINDS website (http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf). Furthermore, the reviewers of applications reviewed by the NINDS Scientific Review Branch are reminded of these issues and asked to pay careful attention to the scientific premise of the proposed projects.

We believe that improving how animal studies are reported will raise awareness of the importance of rigorous study design. Such increased awareness will accelerate both scientific progress and the development of new therapies.

Received 21 August; accepted 10 September 2012.

1. Begley, C. G. & Ellis, L. M. Raise standards for preclinical cancer research. *Nature* **483,** 531–533 (2012).
2. Hess, K. R. Statistical design considerations in animal studies published recently in *Cancer Research*. *Cancer Res.* **71,** 625 (2011).
3. Kilkenny, C. *et al.* Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS ONE* **4,** e7824 (2009).
4. Moher, D., Simera, I., Schulz, K. F., Hoey, J. & Altman, D. G. Helping editors, peer reviewers and authors improve the clarity, completeness and transparency of reporting health research. *BMC Med.* **6,** 13 (2008).
5. Prinz, F., Schlange, T. & Asadullah, K. Believe it or not: how much can we rely on published data on potential drug targets? *Nature Rev. Drug Discov.* **10,** 712 (2011). **The first report that many published studies cannot be reproduced by the pharmaceutical industry.**
6. Sena, E., van der Worp, H. B., Howells, D. & Macleod, M. How can we improve the pre-clinical development of drugs for stroke? *Trends Neurosci.* **30,** 433–439 (2007).
7. Steward, O., Popovich, P. G., Dietrich, W. D. & Kleitman, N. Replication and reproducibility in spinal cord injury research. *Exp. Neurol.* **233,** 597–605 (2012).
8. van der Worp, H. B. & Macleod, M. R. Preclinical studies of human disease: time to take methodological quality seriously. *J. Mol. Cell. Cardiol.* **51,** 449–450 (2011).
9. Hackam, D. G. & Redelmeier, D. A. Translation of research evidence from animals to humans. *J. Am. Med. Assoc.* **296,** 1727–1732 (2006). **A study reporting that a large fraction of high-impact publications in highly reputable journals lack important information related to experimental design.**
10. Macleod, M. R. *et al.* Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. *Stroke* **39,** 2824–2829 (2008). **A study demonstrating that lack of reporting of key methodological parameters is associated with bias.**
11. Bebarta, V., Luyten, D. & Heard, K. Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad. Emerg. Med.* **10,** 684–687 (2003).
12. Crossley, N. A. *et al.* Empirical evidence of bias in the design of experimental stroke studies – A metaepidemiologic approach. *Stroke* **39,** 929–934 (2008).
13. Rooke, E. D., Vesterinen, H. M., Sena, E. S., Egan, K. J. & Macleod, M. R. Dopamine agonists in animal models of Parkinson's disease: a systematic review and meta-analysis. *Parkinsonism Relat. Disord.* **17,** 313–320 (2011).
14. Vesterinen, H. M. *et al.* Improving the translational hit of experimental treatments in multiple sclerosis. *Mult. Scler. J.* **16,** 1044–1055 (2010).
15. Stroke Therapy Academic Industry Roundtable (STAIR). Recommendations for standards regarding preclinical neuroprotective and restorative drug development. *Stroke* **30,** 2752–2758 (1999).
16. Fanelli, D. "Positive" results increase down the hierarchy of the sciences. *PLoS ONE* **5,** e10068 (2010).
17. Jerndal, M. *et al.* A systematic review and meta-analysis of erythropoietin in experimental stroke. *J. Cereb. Blood Flow Metab.* **30,** 961–968 (2010).
18. Macleod, M. R., O'Collins, T., Horky, L. L., Howells, D. W. & Donnan, G. A. Systematic review and metaanalysis of the efficacy of FK506 in experimental stroke. *J. Cereb. Blood Flow Metab.* **25,** 713–721 (2005).
19. Sena, E. S. *et al.* Factors affecting the apparent efficacy and safety of tissue plasminogen activator in thrombotic occlusion models of stroke: systematic review and meta-analysis. *J. Cereb. Blood Flow Metab.* **30,** 1905–1913 (2010).
20. Wheble, P. C. R., Sena, E. S. & Macleod, M. R. A systematic review and meta-analysis of the efficacy of piracetam and piracetam-like compounds in experimental stroke. *Cerebrovasc. Dis.* **25,** 5–11 (2008).
21. Festing, M. F. & Altman, D. G. Guidelines for the design and statistical analysis of experiments using laboratory animals. *ILAR J.* **43,** 244–258 (2002).
22. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8,** e1000412 (2010).
23. van der Worp, H. B. *et al.* Can animal models of disease reliably inform human studies? *PLoS Med.* **7,** e1000245 (2010).
24. Fisher, M. *et al.* Update of the stroke therapy academic industry roundtable preclinical recommendations. *Stroke* **40,** 2244–2250 (2009).
25. Ludolph, A. C. *et al.* Guidelines for preclinical animal research in ALS/MND: a consensus meeting. *Amyotroph. Lateral Scler.* **11,** 38–45 (2010).
26. Shineman, D. W. *et al.* Accelerating drug discovery for Alzheimer's disease: best practices for preclinical animal studies. *Alzheimers Res. Ther.* **3,** 28 (2011).
27. Unger, E. F. All is not well in the world of translational research. *J. Am. Coll. Cardiol.* **50,** 738–740 (2007).
28. Ioannidis, J. P. A. Why most published research findings are false. *PLoS Med.* **2,** e124 (2005).
29. Dienes, Z. Bayesian versus orthodox statistics: which side are you on? *Perspect. Psychol. Sci.* **6,** 274–290 (2011).
30. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* **22,** 1359–1366 (2011).
31. Beal, K. G. & Khamis, H. J. A problem in statistical-analysis: simultaneous inference. *Condor* **93,** 1023–1025 (1991).
32. Lazic, S. E. The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? *BMC Neurosci.* **11,** 5 (2010).
33. Scott, S. *et al.* Design, power, and interpretation of studies in the standard murine model of ALS. *Amyotroph. Lateral Scler.* **9,** 4–15 (2008). **An enlightening analysis of how small sample sizes can lead to false-positive outcomes.**
34. Proschan, M. A. & Waclawiw, M. A. Practical guidelines for multiplicity adjustment in clinical trials. *Control. Clin. Trials* **21,** 527–539 (2000).
35. Festing, M. F. W. Design and statistical methods in studies using animal models of development. *ILAR J.* **47,** 5–14 (2006).
36. Nakagawa, S. & Cuthill, I. C. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* **82,** 591–605 (2007).
37. Chalmers, T. C., Celano, P., Sacks, H. S. & Smith, H. Bias in treatment assignment in controlled clinical-trials. *N. Engl. J. Med.* **309,** 1358–1361 (1983).
38. Jüni, P., Altman, D. G. & Egger, M. Systematic reviews in health care - assessing the quality of controlled clinical trials. *Br. Med. J.* **323,** 42 (2001).
39. Pildal, J. *et al.* Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int. J. Epidemiol.* **36,** 847–857 (2007).
40. Pocock, S. J., Hughes, M. D. & Lee, R. J. Statistical problems in the reporting of clinical-trials. A survey of three medical journals. *N. Engl. J. Med.* **317,** 426–432 (1987).
41. Schulz, K. F., Chalmers, I., Hayes, R. J. & Altman, D. G. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *J. Am. Med. Assoc.* **273,** 408–412 (1995).
42. Wood, L. *et al.* Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *Br. Med. J.* **336,** 601–605 (2008).
43. Moher, D. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *Br. Med. J.* **340,** c869 (2011).
44. Moher, D., Schulz, K. F. & Altman, D. G. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* **357,** 1191–1194 (2001). **Revision of guidelines by the CONSORT group to improve the reporting of randomized clinical trials.**
45. Schulz, K. F., Altman, D. G. & Moher, D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med.* **7,** e1000251 (2010).
46. Plint, A. C. *et al.* Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med. J. Aust.* **185,** 263–267 (2006).
47. Kane, R. L., Wang, J. & Garrard, J. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. *J. Clin. Epidemiol.* **60,** 241–249 (2007).
48. Prady, S. L., Richmond, S. J., Morton, V. M. & Macpherson, H. A systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. *PLoS ONE* **3,** e1577 (2008).
49. Smith, B. A. *et al.* Quality of reporting randomized controlled trials (RCTs) in nursing literature: application of the consolidated standards reporting trials (CONSORT). *Nurs. Outlook* **56,** 31–37 (2008).
50. Macleod, M. R., O'Collins, T., Howells, D. W. & Donnan, G. A. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* **35,** 1203–1208 (2004).

51. Macleod, M. R., O'Collins, T., Horky, L. L., Howells, D. W. & Donnan, G. A. Systematic review and meta-analysis of the efficacy of melatonin in experimental stroke. *J. Pineal Res.* **38,** 35–41 (2005).
52. Gallo, J. M. Pharmacokinetic/pharmacodynamic-driven drug development. *Mt Sinai J. Med.* **77,** 381–388 (2010).
53. Moher, D. *et al.* Describing reporting guidelines for health research: a systematic review. *J. Clin. Epidemiol.* **64,** 718–742 (2011).
54. Callaham, M. L., Wears, R. L., Weber, E. J., Barton, C. & Young, G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *J. Am. Med. Assoc.* **280,** 254–257 (1998).
55. Dickersin, K. & Chalmers, I. Recognizing, investigation and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the WHO. *J. R. Soc. Med.* **104,** 532–538 (2011).
56. Fanelli, D. Negative results are disappearing from most disciplines and countries. *Scientometrics* **90,** 891–904 (2012).
57. Kyzas, P. A., Denaxa-Kyza, D. & Ioannidis, J. P. A. Almost all articles on cancer prognostic markers report statistically significant results. *Eur. J. Cancer* **43,** 2559–2579 (2007).
58. Liu, S. Dealing with publication bias in translational stroke research. *J. Exp. Stroke Transl. Med.* **2,** 16–21 (2009).
59. Rockwell, S., Kimler, B. E. & Moulder, J. E. Publishing negative results: the problem of publication bias. *Radiat. Res.* **165,** 623–625 (2006).
60. Rosenthal, R. The file drawer problem and tolerance for null results. *Psychol. Bull.* **86,** 638–641 (1979).
61. Sterling, T. D. Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *J. Am. Stat. Assoc.* **54,** 30–34 (1959).
62. Song, F. *et al.* Dissemination and publication of research findings: an updated review of related biases. *Health Technol. Assess.* **14,** 1–220 (2010).
63. Sena, E. S., van der Worp, H. B., Bath, P. M. W., Howells, D. W. & Macleod, M. R. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol.* **8,** e1000344 (2010).
64. Fanelli, D. Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE* **5,** e10271 (2010).

Acknowledgements Funded by NINDS.

Author Contributions R.F., A.K.G., S.C.L., J.D.P., S.D.S., U.U. and W.K. organized the workshop. R.B.D., S.E.L., S.C.L., M.R.M. and S.D.S. wrote the manuscript. All authors participated in the workshop and contributed to the editing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.D.S. (silberbs@ninds.nih.gov).

doi:10.1038/nature11458

A roadmap for graphene

K. S. Novoselov[1], V. I. Fal'ko[2], L. Colombo[3], P. R. Gellert[4], M. G. Schwab[5] & K. Kim[6]

Recent years have witnessed many breakthroughs in research on graphene (the first two-dimensional atomic crystal) as well as a significant advance in the mass production of this material. This one-atom-thick fabric of carbon uniquely combines extreme mechanical strength, exceptionally high electronic and thermal conductivities, impermeability to gases, as well as many other supreme properties, all of which make it highly attractive for numerous applications. Here we review recent progress in graphene research and in the development of production methods, and critically analyse the feasibility of various graphene applications.

Could graphene become the next disruptive technology, replacing some of the currently used materials and leading to new markets? Is it versatile enough to revolutionize many aspects of our life simultaneously? In terms of its properties, graphene certainly has the potential. Graphene is the first two-dimensional (2D) atomic crystal available to us. A large number of its material parameters—such as mechanical stiffness, strength and elasticity, very high electrical and thermal conductivity, and many others[1,2]—are supreme. These properties suggest that graphene could replace other materials in existing applications. However, that all these extreme properties are combined in one material means that graphene could also enable several disruptive technologies. The combination of transparency, conductivity and elasticity will find use in flexible electronics, whereas transparency, impermeability and conductivity will find application in transparent protective coatings and barrier films; and the list of such combinations is continuously growing. However, is graphene special and versatile enough to justify the inconveniences of switching to a new technology, usually a lengthy and expensive process?

Graphene properties

One reason that graphene research has progressed so fast is that the laboratory procedures enabling us to obtain high-quality graphene are relatively simple and cheap. Many graphene characteristics measured in experiments have exceeded those obtained in any other material, with some reaching theoretically predicted limits: room-temperature electron mobility of $2.5 \times 10^5\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$ (ref. 3) (theoretical limit[4] $\sim 2 \times 10^5\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$); a Young's modulus of 1 TPa and intrinsic strength of 130 GPa (ref. 5, very close to that predicted by theory[6]); very high thermal conductivity (above 3,000 $\mathrm{W\,mK^{-1}}$; ref. 7); optical absorption of exactly $\pi\alpha \approx 2.3\%$ (in the infrared limit, where α is the fine structure constant)[8]; complete impermeability to any gases[9], ability to sustain extremely high densities of electric current (a million times higher than copper)[10]. Another property of graphene, already demonstrated[11–13], is that it can be readily chemically functionalized.

Graphene's many superior properties justify its nickname of a 'miracle material'. However, some of these characteristics have been achieved only for the highest-quality samples (mechanically exfoliated graphene[14]) and for graphene deposited on special substrates like hexagonal boron nitride[3,15]. As yet, equivalent characteristics have not been observed on graphene prepared using other techniques, although these methods are rapidly improving. Graphene will be of even greater interest for industrial applications when mass-produced graphene has the same outstanding performance as the best samples obtained in research laboratories.

Nature provides us with many other 2D crystals, such as boron nitride and molybdenum disulphide[16]. Being structurally related to graphene but having their own distinctive properties, they offer the possibility of fine-tuning material and device characteristics to suit a particular technology better or to be used in combination with graphene (for example, 2D-based heterostructures[17,18]). Being part of such a large and diverse family of 2D crystals and heterostructures will improve graphene's chances of commercial success, although we do not cover these other 2D crystals in this Review (see Box 1).

Challenges in production

The market of graphene applications is essentially driven by progress in the production of graphene with properties appropriate for the specific application, and this situation is likely to continue for the next decade or at least until each of graphene's many potential applications meets its own requirements. Currently, there are probably a dozen methods being used and developed to prepare graphene of various dimensions, shapes and quality. Here we concentrate only on those that are scalable.

It is logical to categorize these by the quality of the resulting graphene (and thus the possible applications): (1) graphene or reduced graphene oxide flakes for composite materials, conductive paints, and so on; (2) planar graphene for lower-performance active and non-active devices; and (3) planar graphene for high-performance electronic devices. The properties of a particular grade of graphene (and hence the pool of applications that can utilize it) depend very much on the quality of the material, type of defects, substrate, and so forth, which are strongly affected by the production method; see Fig. 1 and Table 1.

Liquid phase and thermal exfoliation

Liquid-phase exfoliation of graphite[19,20] (or any other layered material[21]) is based on exposing the materials to a solvent with a surface tension that favours an increase in the total area of graphite crystallites. The solvent is typically non-aqueous, but aqueous solutions with surfactant can also be used. With the aid of sonication, graphite splits into individual platelets, and prolonged treatment yields a significant fraction of monolayer flakes in the suspension, which can be further enriched by centrifugation.

A related method is the graphite oxide route in which graphite pellets are first oxidized and then ultrasonically exfoliated in an aqueous solution[22]. After exfoliation of graphite oxide the suspension may be further processed by centrifugation, and can then be deposited as a thin film on almost any surface and reduced (albeit partially) *in situ* back to the parent graphene state.

---

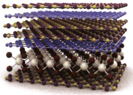
[1]School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK. [2]Department of Physics, Lancaster University, Lancaster LA1 4YB, UK. [3]Texas Instruments Incorporated, 13121 TI Boulevard, MS-365 Dallas, Texas 75243, USA. [4]AstraZeneca, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK. [5]BASF SE, Carbon Materials Innovation Center, Carl-Bosch-Straße 38, 67056 Ludwigshafen, Germany. [6]Samsung Advanced Institute of Technology (SAIT), Samsung Electronics, Yongin-Si, Gyeonggi-Do 446-712, South Korea.

BOX 1

Off-road with other 2D atomic crystals and their heterostructures

The study of graphene has triggered experiments on many other 2D atomic crystals, such as BN, $NbSe_2$, TaS_2, MoS_2 and many others. Similar strategies to those applied to graphene can be used to obtain new 2D materials by mechanical[16] or liquid-phase exfoliation of layered materials[21] or CVD growth. Another way to create new 2D crystals is to start with an existing one (like graphene) and use it as an atomic scaffolding, modifying it by chemical means (graphane[11] and fluorographene[13] are good examples). The pool of possible 2D crystals is huge, covering a massive range of properties: from the most insulating to the most conductive, from the strongest to the softest.

If 2D materials provide a large range of different properties, sandwich structures (made up of two, three, four or more different layers of such materials) can offer even greater scope. These 2D-based heterostructures[17,18] can be tailored with atomic precision and individual layers of very different character can be combined, so the properties of these structures can be tuned to fit an enormous range of possible applications. Furthermore, the functionality of heterostructure stacks is 'embedded' in their design. The first examples have already started to appear, with vertical tunnelling transistors based on this type of heterostructures having been demonstrated recently, showing very promising characteristics[51].



Box 1 Figure | Example of optically active 2D-based heterostructure. Two graphene layers are separated by several layers of boron nitride, which serve as a tunnelling barrier. A built-in electric field (created by the proximity of one of the graphene layers to a monolayer of MoS_2) separates the electron–hole pair, which is created by an incoming photon, resulting in a photocurrent.

An industrially important variation of the fully aqueous-based graphite-oxide route makes use of a thermal-shock procedure to achieve exfoliation and reduction simultaneously[23]. Even though the resulting material may contain graphene components with several layers, it still preserves many of the appealing properties of single-layer graphene. Similarly to oxidation, the parent graphite stacking can be disturbed via intercalation of small molecules. Such graphite intercalation compounds may then be used in a similar way as precursors and can subsequently be subjected to thermal or plasma processes to achieve their delamination into single sheets.

Also, there are several methods of producing suspensions of graphene nanoribbons—via unzipping of single-wall carbon nanotubes[24,25]. Although they are more expensive than chemical exfoliation of graphite or graphite oxide, these methods allow one to achieve suspensions with well-defined distributions (potentially very narrow) of graphene platelets. Similarly, nanotube unzipping allows better control over the chemical functionalization and quality of the edges.

Such bulk grades of graphene are already available on the tonne scale and are currently being evaluated in numerous fields of application[26]. Thus, graphene-based paints and inks will find their way into printed electronics, electromagnetic shielding, barrier coatings, heat dissipation, supercapacitors, smart windows[27], and so on. A number of flake-based products can be expected in the marketplace within a few years, and prototype applications for conductive inks have already been demonstrated on the commercial level.

Chemical vapour deposition

Large-area uniform polycrystalline graphene films are now being grown by chemical vapour deposition (CVD) on copper foils and films, and show promise for many applications[28]. Despite the fact that the complete process typically requires transfer from the copper support to a dielectric surface or other substrate of interest[1], the production of square metres of graphene has already been achieved[29]. These films have also been transferred onto 200-mm Si wafers on which state-of-the-art devices have been demonstrated. On a smaller scale, these films show transport properties equivalent to those of exfoliated graphene on both SiO_2 and hexagonal boron nitride substrates. Despite the presence of defects, grain boundaries, inclusions of thicker layers, and so on, such films are ready for use in transparent conductive coating applications (such as touch screens).

At present, the process is expensive owing to large energy consumption and because the underlying metal layer has to be removed. However, once the transfer process is optimized this method may indeed be disruptive and cost-effective. A number of issues need to be resolved before graphene CVD technology can become widely used. Graphene growth on thin (tens of nanometres) films of metals needs to be achieved, simultaneously gaining control of the domain (grain) size, ripples, doping level and the number of layers. Control of the number and relative crystallographic orientation of the graphene layers is critical because it will enable a number of applications which would require double, triple and even thicker layers of graphene. Simultaneously, the transfer process should be improved and optimized with the objectives of minimizing the damage to graphene and of recovering the sacrificial metal.

The transfer process might be as complicated as the growth of graphene itself. However, there are a number of applications which rely on conformal growth of graphene on the surface of the metal, and do not require graphene transfer at all: high thermal and electrical conductivities as well as excellent barrier properties allow graphene greatly to enhance the performance of copper interconnects in integrated circuits. Also, because graphene is inert, it is an excellent barrier for any gas, and it forms a conformal layer on metal surfaces with the most complex topographies: such coatings can protect against corrosion.

The game-changing breakthroughs would be the development of graphene growth on arbitrary surfaces and/or at low temperatures (for example, using plasma-enhanced CVD or other methods) with a minimal number of defects. The former would allow one to avoid the complex and expensive transfer step and promote better integration of this 2D crystal with other materials (like Si or GaAs). The latter would improve compatibility with modern microelectronic technologies and allow significant energy saving.

Synthesis on SiC

Silicon carbide is a common material used for high-power electronics. It has been demonstrated that graphitic layers can be grown either on the silicon or carbon faces of a SiC wafer by sublimating Si atoms, thus leaving a graphitized surface[30]. Initially, the C-terminated face of SiC was used to grow a turbostratic stack of many randomly oriented polycrystalline layers[31], but now the number of graphene layers grown[32] can be controlled. The quality of such graphene can be very high, with crystallites approaching hundreds of micrometres in size[33].

The two major drawbacks of this method are the high cost of the SiC wafers and the high temperatures (above 1,000 °C) used, which are not directly compatible with silicon electronics technology. There are potentially several ways to take advantage of the growth of graphene on SiC, including the growth of thin SiC on Si, although this approach requires further development. As a result of the high-temperature growth, high substrate cost, and small-diameter wafers, the use of graphene on SiC

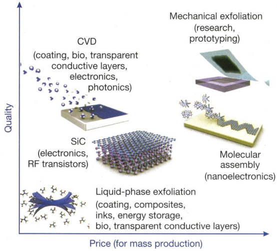

Mechanical exfoliation
(research, prototyping)
CVD
(coating, bio, transparent conductive layers, electronics, photonics)
Quality
SiC
(electronics, RF transistors)
Molecular assembly
(nanoelectronics)
Liquid-phase exfoliation
(coating, composites, inks, energy storage, bio, transparent conductive layers)
Price (for mass production)

Figure 1 | There are several methods of mass-production of graphene, which allow a wide choice in terms of size, quality and price for any particular application.

will probably be limited to niche applications. High-frequency transistors based on SiC-grown graphene[34] may well find applications within a decade when the existing technology, based on III–V materials (such as InGaAs, GaN, and so on) reaches its limit at about 1 THz. The short gate transistors that are currently widely used make even the 20-μm size domains (currently achieved in graphene grown on SiC) suitable for such applications. Another very attractive, though niche, application of this type of graphene is in metrological resistance standards[35], where samples of graphene grown on SiC have already been demonstrated to deliver higher resistance accuracy at higher temperatures than do conventionally used GaAs heterostructures.

Apart from the high temperature required for growth, which currently seems to be an insurmountable problem, the other issues that need to be addressed in the next decade are the elimination of terraces, the growth of the second or third layers at the edges of the terraces (which also strongly contribute to carrier scattering), an increase in the size of the crystallites and control of unintentional doping from the substrate and buffer layers.

Other growth methods

Although there are a number of other growth methods, it is unlikely that they will become commercially viable in the next decade. Nevertheless, some of these methods have certain advantages and should be researched further. Surface-assisted coupling of molecular monomer precursors into linear polyphenylenes with subsequent cyclodehydrogenation is an exciting way to create high-quality graphene nanoribbons and even more complex structures (like T- and Y-shaped connections)[36] using a chemistry-driven bottom-up approach. Molecular beam epitaxy has been used to grow chemically pure graphene[37], but it is unlikely to be used on a large scale because of its much higher cost than CVD methods. Laser ablation is a potentially interesting growth technique allowing the deposition of graphene nanoplatelets on arbitrary surfaces[38]. This relatively expensive method is in direct competition with the spray-coating of chemically exfoliated graphene, so it is unlikely to be widely used.

Graphene electronics

It is unlikely that graphene will make it into high-performance integrated logic circuits as a planar channel material within the next decade because of the absence of a bandgap. However, many other, less stringent, graphene electronic applications are being developed, using the available (probably not ideal in terms of quality) material. Figure 2 and Table 2 list some of the possible applications and the time that it may take for graphene-based prototypes to be demonstrated.

Flexible electronics

Transparent conductive coatings are widely used in electronic products such as touch screen displays, e-paper (electronic paper) and organic light-emitting diodes (OLEDs) and require a low sheet resistance with high transmittance (of over 90%) depending on the specific application. Graphene meets the electrical and optical requirements (sheet resistance reaching 30 Ω per square of 2D area in highly doped samples) and an excellent transmittance of 97.7% per layer[8], although the traditionally used indium tin oxide (ITO) still demonstrates slightly better characteristics. However, considering that the quality of graphene improves every year (already making the difference in performance marginal), while ITO will become more expensive and ITO deposition is already expensive, graphene has a chance of securing a good fraction of the market. Graphene also has outstanding mechanical flexibility and chemical durability—very important characteristics for flexible electronic devices[29], in which ITO usually fails.

The requirements of electrical properties (for example, sheet resistance) for each electrode type differ from application to application. Depending on the production methods, various grades of transparent conductive coating could be produced from graphene. Thus, electrodes for touch screens (although requiring an expensive CVD method of production) tolerate a relatively high sheet resistance (50–300 Ω per square) for a transmittance of 90%. The advantage of graphene electrodes in touch panels is that graphene's endurance far exceeds that of any other available candidate at the moment. Moreover, the fracture strain of graphene is ten times higher[5] than that of ITO, meaning that it could also successfully be applied to bendable and rollable devices.

Rollable e-paper is a very appealing electronic product. It requires a bending radius of 5–10 mm, which is easily achievable by a graphene electrode. In addition, graphene's uniform absorption across the visible spectrum[8] is beneficial for colour e-papers. However, the contact resistance between the graphene electrode and the metal line of the driving circuitry is still a problem. A working prototype is expected by 2015, but the manufacturing cost needs to decrease before it will appear on the market.

OLED devices have become an attractive technology and the first (non-graphene) products are expected on the market by 2013. Besides

Table 1 | Properties of graphene obtained by different methods

Method	Crystallite size (μm)	Sample size (mm)	Charge carrier mobility (at ambient temperature) (cm^2 V^{-1} s^{-1})	Applications
Mechanical exfoliation	>1,000	>1	>2 × 10^5 and > 10^6 (at low temperature)	Research
Chemical exfoliation	≤0.1	Infinite as a layer of overlapping flakes	100 (for a layer of overlapping flakes)	Coatings, paint/ink, composites, transparent conductive layers, energy storage, bioapplications
Chemical exfoliation via graphene oxide	~100	Infinite as a layer of overlapping flakes	1 (for a layer of overlapping flakes)	Coatings, paint/ink, composites, transparent conductive layers, energy storage, bioapplications
CVD	1,000	~1,000	10,000	Photonics, nanoelectronics, transparent conductive layers, sensors, bioapplications
SiC	50	100	10,000	High-frequency transistors and other electronic devices

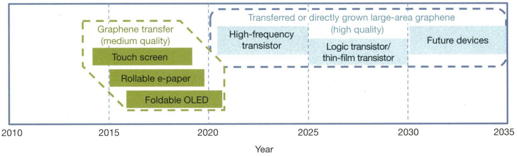


Figure 2 | Graphene-based display and electronic devices. Display applications are shown in green; electronic applications are shown in blue. Possible application timeline, based on projections of products requiring advanced materials such as graphene. The figure gives an indication of when a functional device prototype could be expected based on device roadmaps and the development schedules of industry leaders.

strict requirements for the sheet resistance (below 30 Ω per square), other crucial parameters for such devices are the work function and the electrode's surface roughness, which effectively governs the performance. The tunability of graphene's work function could improve the efficiency, and its atomically flat surface would help avoid electrical shorts and leakage currents. Graphene electrodes have already been demonstrated in OLED test cells[39]. Advanced flexible or foldable OLED devices could be introduced after 2016 once device integration issues (such as conformal deposition of graphene on three-dimensional structures and contact resistance between graphene and the source/drain) are resolved.

In the low-cost sector everything is set up for mass production. Liquid-phase exfoliation produces such graphene coatings without the use of expensive vacuum technology. Although the resistance of these films is on the high side, they still perform well enough for smart windows, solar cells and some touch screen applications. Graphene has the important advantage of flexibility and mechanical strength, which ensures that graphene-based devices will probably dominate flexible applications.

High-frequency transistors

Graphene has been considered and intensively researched for high-frequency transistor applications[34]. However, it has to compete against more mature technologies such as compound semiconductors (III–V materials). Thus, graphene will probably be used only after 2021, when even III–V materials will fail to satisfy device requirements. Projections show that III–V materials will no longer be able to obtain the required cut-off frequency $f_T = 850$ GHz (the top frequency for current modulation) and maximum oscillation frequency $f_{max} = 1.2$ THz (the top frequency for power modulation) after 2021 because device requirements will become more stringent. A recent graphene progress report[40] presented a value of f_T as high as 300 GHz, with the possibility of extending it up to 1 THz at a channel length of about 100 nm (ref. 41). On the other hand, f_{max} has only reached 30 GHz in traditional graphene structures, which is far from the 330 GHz Si high-frequency transistor performance, according to the 2011 International Technology Roadmap for Semiconductors (ITRS). Thus, the principal remaining research issue is the low value of f_{max} for graphene transistors, which trails f_T by an order of magnitude in a comparable conventional device. There are two ways to improve f_{max}: by lowering the gate resistance or the source–drain conductance at pinch-off (ref. 42). The former approach could be done using well-established semiconductor processes. The latter will require current saturation in the graphene high-frequency transistor, which will probably involve finding a new dielectric layer with properties similar to those of boron nitride[43] and compatible with modern semiconductor technology. An f_{max} of 58 GHz has been reported[44] using graphene on top of an exfoliated hexagonal boron nitride film[3,15].

Logic transistor

It is widely accepted that Si technology will be extended to nearly or even below the 10-nm level. Graphene transistors might have an opportunity to replace the silicon technology only after 2020 (according to the 2011 ITRS).

Several research paths are being targeted at opening a bandgap in graphene: nanoribbon[36,45] and single electron transistor[46,47] formation, bilayer control[32,48] and chemically modified graphene[11,13]. However, all of these approaches (apart from chemical modification) have so far been unable to open a bandgap wider than 360 meV (ref. 49), which limits the on/off ratio to about 10^3, much less than the required 10^6. Even worse, they also lead to the degradation of the carrier mobility in graphene[50].

The issue of the low on/off ratio is resolved in the new transistor designs, which exploits the modulation of the work function of graphene, gaining control over vertical (rather than planar) transport through various barriers[51]. Although such devices allow for spectacular on/off ratios of $>10^6$, more work on integration is required to enable the use of graphene for logic applications after 2025.

Graphene's electrical and thermal conductivities as well as its excellent barrier properties might push this material towards being used as interconnects as well as for thermal dissipation in integrated circuits. Graphene can easily be grown on copper by CVD, so we might see this combination used for such applications.

Table 2 | Electronics applications of graphene

Application	Drivers	Issues to be addressed
Touch screen	Graphene has better endurance than benchmark materials	Requires better control of contact resistance, and the sheet resistance needs to be reduced (possibly by doping)
E-paper	High transmittance of monolayer graphene could provide visibility	Requires better control of contact resistance
Foldable OLED	Graphene of high electronic quality has a bendability of below 5 mm, improved efficiency due to graphene's work function tunability, and the atomically flat surface of graphene helps to avoid electrical shorts and leakage current	Requires better control of contact resistance, the sheet resistance needs to be reduced, and conformal coverage of three-dimensional structures is needed
High-frequency transistor	No manufacturable solution for InP high-electron-mobility transistor (low noise) after 2021, according to the 2011 ITRS	Need to achieve current saturation, and $f_T = 850$ GHz, $f_{max} = 1{,}200$ GHz should be achieved
Logic transistor	High mobility	New structures need to resolve the bandgap–mobility trade-off and an on/off ratio larger than 10^6 needs to be achieved

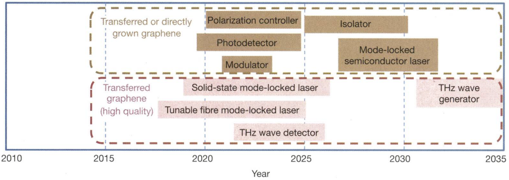


Figure 3 | Graphene-based photonics applications. Optical applications are shown in pink; optical interconnect applications are shown in brown. Possible application timeline, enabled by continued advances in graphene technologies, based on projections of products requiring advanced materials such as graphene. The figure gives an indication of when a functional device prototype could be expected based on device roadmaps and the development schedules of industry leaders.

Photonics

Electrons in graphene behave as massless two-dimensional particles, which leads to a significant wavelength-independent absorption ($\pi\alpha = 2.3\%$) for normal incident light[8] below about 3 eV. Additionally, mono- and bi-layer graphene become completely transparent when the optical energy is smaller than double the Fermi level, owing to Pauli blocking[52]. These properties would suit many controllable photonic devices (Fig. 3 and Table 3).

Photodetectors

Graphene photodetectors are presently one of the most actively studied photonic devices. Unlike semiconductor photodetectors, which have limited detecting spectral width, graphene can in principle be used for a wide spectral range from ultraviolet to infrared. Another advantage of graphene is its high operating bandwidth, which makes it suitable for high-speed data communications. The maximum bandwidths of InGaAs (for optical communication) and Ge (for optical interconnection) photodetectors are limited to 150 GHz (ref. 53) and 80 GHz (ref. 54) respectively, owing to the carrier transit times. The high carrier mobility of graphene enables ultrafast extraction of photo-generated carriers, possibly allowing extremely high bandwidth operation. The transit-time-limited bandwidth of graphene photodetectors is calculated[55] to be 1.5 THz at the reported saturation carrier velocity[56]. In practice, the maximum bandwidth of a graphene photodetector would be limited[55] to 640 GHz by the time constant resulting from the capacitive (*RC*) delay, rather than the transit time.

Owing to the absence of a bandgap, the graphene photodetector requires a different carrier extraction model from that of semiconductor photodetectors. Currently, graphene photodetectors use the local potential variation near the metal–graphene interfaces to extract the photo-generated carriers[57]. Photo-responses of up to 40 GHz (ref. 55) and 10 GHz (ref. 58) detector operation have been demonstrated. However, the maximum responsivity is low (a few $\mathrm{mA\,W^{-1}}$; ref. 58) in comparison to the required $\sim 1\,\mathrm{A\,W^{-1}}$) because of the limited absorption caused by the small effective detection areas and the thinness of graphene.

There are several possible ways to improve the sensitivity of graphene photodetectors, such as by using plasmonic nanostructures for the enhancement of the local optical electric field[59] or by integrating it with a waveguide to increase the light–graphene interaction length[49]. Given the maximum bandwidth of the Ge photodetector and optical interconnection roadmap, a graphene photodetector with a bandwidth over 100 GHz will be competitive after 2020, providing that a method compatible with modern semiconductor technology of growing high-quality graphene (with mobility $>$20,000 $\mathrm{cm^2\,V^{-1}\,s^{-1}}$) is secured.

Optical modulator

Optical modulators are one of the key active building blocks for optical interconnects used to encode transmission data by altering the properties of light[60] such as phase, amplitude, and polarization using electro-refraction or electro-absorption. Si optical modulators, such as Mach–Zehnder interferometers[61], ring resonators[62] and electro-absorption modulators[63] are based on interference, resonance and bandgap absorption, respectively. Their operating spectra are usually narrow, however, and their slow switching times limit operation bandwidths. For Si waveguide modulators, a large resistance in the p–n junction through the Si

Table 3 | Photonic applications of graphene

Application	Drivers	Issues to be addressed
Tunable fibre mode-locked laser	Graphene's wide spectral range	Requires a cost-effective graphene-transferring technology
Solid-state mode-locked laser	Graphene-saturable absorber would be cheaper and easy to integrate into the laser system	Requires a cost-effective graphene-transferring technology
Photodetector	Graphene can supply bandwidth per wavelength of 640 GHz for chip-to-chip or intrachip communications (not possible with IV or III–V detectors)	Need to increase responsivity, which might require a new structure and/or doping control, and the modulator bandwidth must follow suit
Polarization controller	Current polarization controlling devices are bulky or difficult to integrate but graphene is compact and easy to integrate with Si	Need to gain full control of parameters of high-quality graphene
Optical modulator	Graphene could increase operating speed (Si operation bandwidth is currently limited to about 50 GHz), thus avoiding the use of complicated III–V epitaxial growth or bonding on Si	High-quality graphene with low sheet resistance is needed to increase bandwidth to over 100 GHz
Isolator	Graphene can provide both integrated and compact isolators on a Si substrate, dramatically aiding miniaturization	Decreasing magnetic field strength and optimization of process architecture are important for the products
Passively mode-locked semiconductor laser	Core-to-core and core-to-memory bandwidth increase requires a dense wavelength-division-multiplexing optical interconnect (which a graphene-saturable absorber can provide) with over 50 wavelengths, not achievable with a laser array	Competing technologies are actively mode-locked semiconductor lasers or external mode-lock lasers but the graphene market will open in the 2020s; however, interconnect architecture needs to consume low power

core regions is a problem, confining bandwidths to usually less than about 50 GHz.

Excellent optical modulator performance can be achieved by exploiting graphene's ability to absorb a small amount of incident light over ultrawide ranges of wavelengths and its ultrafast response. To do this, the interband transitions of photo-generated electrons in a single graphene layer[64] are modulated over broad spectral ranges by a drive voltage, leading to operating speeds with bandwidth exceeding 1 GHz in the near infrared range[65]. With some structural changes, an even wider operation bandwidth of more than 50 GHz has been suggested, using inter-gated dual graphene layers[49] to reduce the resistance in the *RC* delay time, offering a pathway to a regime of hundreds of gigahertz, although such developments are not expected before 2020. Graphene is also promising for THz-range wireless communications[66] where optical losses are an order of magnitude smaller than those in noble metals.

Mode-locked laser/THz generator

Ultrafast passively mode-locked lasers have been used for various applications in spectroscopy, material micromachining[67], bio-medicine[68] and in security applications; they usually use saturable absorbers to cause intensity modulation by selectively transmitting high-intensity light only. Compared with the widely used semiconductor saturable absorbers[69], graphene absorbs a significant amount of photons per unit thickness[8] and therefore reaches saturation at a lower intensity over a wide spectral range[70,71]. Ultrafast carrier relaxation time, controllable modulation depth, high damage threshold, high thermal conductivity[8] and wide spectral range tunability[72] are other benefits of graphene-saturable absorbers[73]. These applications need only a small area of graphene, so commercialization could take place even before 2020.

Most studies are focused on fibre and solid-state lasers[74], but graphene-saturable absorbers can also find applications in semiconductor laser technology. Optical interconnection with a wavelength-division-multiplexing scheme requires a laser array with different wavelengths. One way to provide many different wavelengths is to use a single laser with multiple longitudinal modes, such as a mode-locked laser[75]. An actively mode-locked Si hybrid laser has been studied for this purpose[76], but a graphene-saturable absorber could enable a passively mode-locked semiconductor laser with simple fabrication and operation. However, we expect this application will be useful only after developing a highly integrated optical interconnection around the late 2020s.

THz generators can be used in various applications such as medical imaging, chemical sensors, and security devices. Early proposals based on THz electromagnetic wave generation use graphene as a gain medium to generate stimulated emission by optical pumping[77]. However, electrons and holes have similar mobility values, so the photo-Dember effect (formation of a dipole and resulting THz emission due to the difference in diffusion times of electrons and holes) may not be effective. Hence, it is difficult to obtain a continuous-wave operation overcoming stimulated emission thresholds without damaging the material. Recent studies on THz wave generation suggest using a pulsed excitation of single-layer graphene or using multilayer graphite[78] under a femtosecond laser pulse field to generate carriers that will be accelerated to generate the THz wave. However, the intensity is 10^3–10^4 times weaker than that generated from a III–V-semiconductor-based photoconductive antenna or resonant tunnelling devices[79]. Practical THz-wave generators using graphene are unlikely to emerge before 2030.

Optical polarization controller

Polarization controllers such as polarizers and polarization rotators are crucial passive components with which to manipulate the polarization properties of photons. The differential attenuation of the transverse magnetic mode due to the excitation of Dirac fermions can provide an excellent extinction ratio of 27 dB, covering very broad communication bands. Compact optical polarizers have been demonstrated in data-communication optical fibres integrated with graphene as an in-line conductive layer[80]. High-quality millimetre-sized graphene needs to be integrated with an optical fibre or silicon in a hybrid device. Therefore, if the graphene-processing technology matures, these devices could come into play as early as 2020.

Faraday rotation is a popular way to manipulate light polarization[81]. Landau quantization in the two-dimensional electron gas in graphene[82] results in a giant rotation with a fast response and a broadband tunability. Even larger polarization rotations can be achieved with multi-stacking graphene structures. Two polarizers combined with these Faraday rotators could be made into very compact hybrid isolators. However, a desirable magnetic field smaller than 1 T will be a serious challenge for graphene isolators, delaying its debut until the late 2020s.

Composite materials, paints and coating

Graphene-based paints can be used for conductive ink, antistatic, electromagnetic-interference shielding, and gas barrier applications. In principle, the production technology is simple and reasonably developed, with most of the major graphite mining companies as well as new start-up companies having programmes on liquid-phase or thermally exfoliated graphene. In addition, over the next few years chemical derivatives of graphene will be heavily developed to control the conductivity and optical opacity of the products.

Graphene is highly inert, and so can also act as a corrosion barrier against water and oxygen diffusion. Given that it can be grown directly on the surface of almost any metal under the right conditions, it could form a protective conformal layer, that is, it could be used on complex surfaces.

The mechanical, chemical, electronic and barrier properties of graphene along with its high aspect ratio make graphene attractive for applications in composite materials. The commercial position held by carbon fibres, however, is so strong that graphene will need substantial development before it will be economically feasible to use it as the main reinforcement component. The target is to achieve a 250 GPa Young's modulus at the price of €25 per kilogram. In addition, pure graphene might not have the same adhesion properties to the matrix as carbon fibres, which would require more chemical modification of graphene.

An equally large market exists in bringing extra functionality to composites, where the scope of graphene might be large and possibly realized more rapidly. Graphene can contribute gas and moisture barrier properties, electromagnetic shielding, electrical and thermal conductivity, and a strain monitoring capability to the surrounding polymer matrix. As an additive to a composite matrix polymer it might increase the operating temperature level of composites, reduce moisture uptake, induce antistatic behaviour, give lightning strike protection and improve composite compressive strength. There are also a number of applications for which it is difficult to use carbon fibres that would still benefit from excellent mechanical reinforcement (injection-moulded composites).

Considering that many companies involved in the carbon business have already established programs on graphene and graphene-oxide production, it is possible to expect graphene-based composites to appear on the market within a few years. The real breakthrough, however, will be expected when graphene flakes over 10 μm in size are easily obtainable—the dimension required to use in full the advantage of the high Young's modulus of graphene[5,83]. Fortunately, it has been demonstrated that graphite flakes thicker than one monolayer can provide a significant level of reinforcement[83], thus making the implementation of graphene-based composites realistic within a shorter time.

Energy generation and storage

There is a constant search for highly efficient renewable energy technologies, and it would be surprising if graphene were not involved in this race. At present, most efforts are concentrated on solar cells, which could be divided into those where graphene acts as the active medium and those that use graphene as a transparent or distributed electrode material. The former use the same principle of operation as already discussed for photodetectors, and, in principle, would benefit from uniform absorption over a broad spectrum[8]. However, owing to the low intrinsic optical absorption of graphene[8], such devices would require complex interferometry or plasmonic enhancement structures[59] to achieve the desired responsivity, and thus are unlikely to be widely used soon. Instead, the use of graphene as a transparent electrode in either quantum dots or dye-sensitized solar cells has proved highly beneficial. Doping can vary the position of the Fermi level in graphene significantly, so such electrodes have been used both as electron[84] and hole[85] conducting media. With the cost of graphene produced by liquid-phase or thermal exfoliation going down[26] we can expect wide use of graphene in dye solar cells, especially in applications where mechanical flexibility is paramount.

The use of graphene in next-generation lithium-ion batteries is currently being widely studied. Traditionally used in commercial lithium-ion batteries, cathodes frequently suffer from poor electrical conductivity, which is overcome by the addition of graphite and carbon black to the electrode formulation. Graphene, with its sheet-like morphology, would not only act as an advanced conductive filler but may also give rise to novel core–shell or sandwich-type nanocomposite structures[86]. The resulting increase in electrical conductivity of these new morphologies would help in overcoming one of the key limitations of lithium-ion batteries—their low specific power density. Lastly, the high thermal conductivity of graphene may be advantageous when it comes to high current loads that generate significant amounts of heat within the battery system. As anodes, graphene nanosheets can be used to intercalate lithium reversibly into the layered crystals. Graphene nanosheets used in conjunction with carbon nanotubes and fullerenes, C_{60}, increased the battery charge capacity[87].

Supercapacitors (Fig. 4) are based on storage of energy within electrochemical double-layer capacitors[88]. The superior rate performance of state-of-the-art devices (compared to lithium-ion batteries) is based on predominantly electrostatic storage of electrical energy and is determined by the combination of a high-surface-area activated carbon material and a nanoscopic charge separation at the electrode–electrolyte interface. Graphene is an obvious material choice for this application[1], offering high intrinsic electrical conductivity, an accessible and defined pore structure, good resistance to oxidative processes and high temperature stability. Currently the prototype graphene-based electrochemical double-layer capacitors[89] lead the field in capacitance as well as energy and power densities. Although the characteristics of graphene supercapacitors are very encouraging, there are still issues which must be addressed before the commercial use of such systems. In particular, the irreversible capacitance of graphene-based supercapacitors is still too high, which could probably be improved by reducing the number of defects or choosing a better electrolyte.

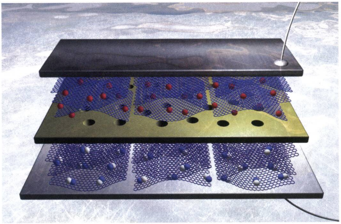


Figure 4 | In a supercapacitor device two high-surface-area graphene-based electrodes (blue and purple hexagonal planes) are separated by a membrane (yellow). Upon charging, anions (white and blue merged spheres) and cations (red spheres) of the electrolyte accumulate at the vicinity of the graphene surface. The ions are electrically isolated from the carbon material by the electrochemical double layer that is serving as a molecular dielectric.

There are also reports on the use of graphene nanosheets as a support material for platinum catalysts for fuel cells[90]. Unlike carbon black, which is the baseline support material for platinum catalysts, graphene decreases the platinum particle size to under a nanometre because of the strong interaction between the platinum atoms and graphene. The strong interaction of platinum and graphene and the small particle size is leading to increased catalytic activity in direct methanol fuel cells.

Common benchmark materials in energy-related applications (graphite, carbon black and activated carbon) will only be replaced if graphene is proved to be superior in terms of both performance and cost. That graphene of suitable grades for such applications is already available in scalable amounts[26] might speed its progression into real devices.

Graphene for sensors and metrology

Graphene, being a two-dimensional fabric and a surface without bulk, has properties that are extremely sensitive to the environment. Thus, it is natural to consider using graphene for sensor applications, from measurements of magnetic field to DNA sequencing and from the monitoring of the velocity of surrounding liquid to strain gauges. The latter (with either electrical or optical readouts) are probably the most competitive application. Graphene is the only crystal which can be stretched by 20%, thus enhancing the working range of such sensors significantly[5].

Currently, graphene gas detectors, although extremely sensitive, have only a minor competitive edge over existing devices. Low selectivity and poisoning by water limit their area of applicability, although such detectors can be produced so cheaply that they could be used in certain niche applications. Functionalization might improve the selectivity of graphene sensors, but because it is rather an expensive method, it is probably most suitable for bio-sensing.

The major advantage of graphene sensors is their multi-functionality. A single device can be used in multidimensional measurements (for example, strain, gas environment, pressure and magnetic field). In this sense graphene offers unique opportunities. With the development of increasingly interactive consumer electronic devices, such sensors will certainly find their way into many products.

The unique bandstructure of graphene, with its anomalously large energy splitting between the zero-energy and the first Landau levels, makes it an ideal material to develop the universal resistance standard based on the quantum Hall effect[1]. The precision of quantum Hall effect quantization of 0.1 parts per billion for epitaxial graphene grown on the Si face of SiC by far outperforms that in the traditionally used GaAs heterostructures[35,91], and such devices are already being used by several metrological facilities.

Bioapplications

Graphene has a number of properties which make it potentially promising for bioapplications. Its large surface area, chemical purity and the possibility of easy functionalization provide opportunities for drug delivery. Its unique mechanical properties suggest tissue-engineering applications and regenerative medicine[92]. Its combination of ultimate thinness, conductivity and strength make it an ideal support for imaging biomolecules in transmission electron microscopy[93]. Also, chemically functionalized graphene might lead to fast and ultrasensitive measurement devices, capable of detecting a range of biological molecules including glucose, cholesterol, haemoglobin and DNA[94].

As a result of their large surface area and delocalized π electrons, graphene derivatives can solubilize and bind drug molecules and thus have the potential to be drug delivery vehicles in their own right if sufficiently high drug loadings and suitable *in vivo* drug distribution and release profiles can be achieved. Graphene is also lipophilic, which might help in solving another challenge in drug delivery—membrane barrier penetration (Fig. 5). Most of the limited work that has been done so far has focused on investigating the loading and *in vitro* behaviour for aromatic anticancer drugs such as doxorubicin[95]. Intravenous

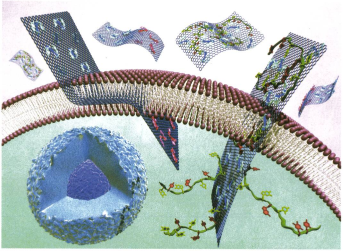


Figure 5 | Manipulating the hydrophilic–lipophilic properties of graphene (blue hexagonal planes) through chemical modification would allow interactions with biological membranes (purple-white double layer), such as drug delivery into the interior of a cell (blue region).

administration of polyethylene glycol-modified graphene oxide, labelled with a near-infrared fluorescence dye but not carrying any drug, has shown passive tumour targeting in mouse xenograft models. The tumours were killed when irradiated with a low-power near-infrared laser, showing the potential of using graphene derivatives for photothermal cancer treatment[96]. However, given the high safety, clinical and regulatory hurdles and long timescales associated with drug development, which are exacerbated when new materials are involved, it is unlikely that products using graphene-based drug delivery technology will be near the market before 2030.

Tissue engineering is an emerging area of technology with potential for a significant impact on patient treatment across a range of disease areas, although as yet only a small number of potential products have entered clinical trials. Graphene could be incorporated into the scaffold materials used for tissue engineering to improve their mechanical (strength and elasticity) and selective barrier[97] properties and potentially to modulate their biological performance in areas such as cell adhesion, proliferation and differentiation[95].

Before graphene can fulfil its promise in the biomedical area we must understand its biodistribution, biocompatibility and acute and chronic toxicity under conditions that are relevant to exposure during manufacture and subsequent use. Ultimately, this will probably need to be done for the particular form of graphene being used in a given application because the outcome is likely to vary with size, morphology and chemical structure. In some cases it may also be possible to exploit the biological activity that gives rise to a particular toxicity profile. For example, a 'toxic' graphene derivative could potentially be therapeutic in its own right as an antibiotic or anticancer treatment.

Conclusions

Physicists are used to thinking of graphene as a perfect two-dimensional lattice of carbon atoms. However, the paradigm is now shifting as pure science opens new technology routes: even less-than-perfect layers of graphene can be used in certain applications. In fact, different applications require different grades of graphene, bringing closer widespread practical implementation of this material.

As the current market for graphene applications is driven by the production of this material, there is a clear hierarchy in how soon the applications will reach the user or consumer. Those that use the lowest-grade, cheapest and most available material will be the first to appear, probably in a few years, and those which require the highest, electronic-quality grades or biocompatibility may well take decades to develop. Also, because developments in the last few years were extremely rapid, graphene's prospects continue to improve. Nevertheless, established benchmark materials will only be replaced if the properties of graphene, however appealing, can be translated into applications that are sufficiently competitive to justify the cost and disruption of changing existing industrial processes.

Graphene is a unique crystal in the sense that it combines many superior properties, from mechanical to electronic. This suggests that its full power will only be realized in novel applications, which are designed specifically with this material in mind, rather than when it is used to replace other materials in existing applications. Interestingly, such an opportunity is likely to be provided very soon with development of such new technologies as printable and flexible electronics, flexible solar cells and supercapacitors.

Received 5 April; accepted 13 July 2012.

1. Geim, A. K. & Novoselov, K. S. The rise of graphene. *Nature Mater.* **6,** 183–191 (2007).
2. Geim, A. K. Graphene: status and prospects. *Science* **324,** 1530–1534 (2009).
3. Mayorov, A. S. *et al.* Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano Lett.* **11,** 2396–2399 (2011).
4. Morozov, S. V. *et al.* Giant intrinsic carrier mobilities in graphene and its bilayer. *Phys. Rev. Lett.* **100,** 016602 (2008).
5. Lee, C., Wei, X. D., Kysar, J. W. & Hone, J. Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science* **321,** 385–388 (2008).
6. Liu, F., Ming, P. M. & Li, J. Ab initio calculation of ideal strength and phonon instability of graphene under tension. *Phys. Rev. B* **76,** 064120 (2007).
7. Balandin, A. A. Thermal properties of graphene and nanostructured carbon materials. *Nature Mater.* **10,** 569–581 (2011).
8. Nair, R. R. *et al.* Fine structure constant defines visual transparency of graphene. *Science* **320,** 1308 (2008).
9. Bunch, J. S. *et al.* Impermeable atomic membranes from graphene sheets. *Nano Lett.* **8,** 2458–2462 (2008).
10. Moser, J., Barreiro, A. & Bachtold, A. Current-induced cleaning of graphene. *Appl. Phys. Lett.* **91,** 163513 (2007).
11. Elias, D. C. *et al.* Control of graphene's properties by reversible hydrogenation: evidence for graphane. *Science* **323,** 610–613 (2009).
12. Loh, K. P., Bao, Q. L., Ang, P. K. & Yang, J. X. The chemistry of graphene. *J. Mater. Chem.* **20,** 2277–2289 (2010).
13. Nair, R. R. *et al.* Fluorographene: a two-dimensional counterpart of Teflon. *Small* **6,** 2877–2884 (2010).
14. Novoselov, K. S. *et al.* Electric field effect in atomically thin carbon films. *Science* **306,** 666–669 (2004).
    In this paper a micromechanical cleavage method was used to obtain high-quality sheets of graphene and its transport and switching properties were studied.
15. Dean, C. R. *et al.* Boron nitride substrates for high-quality graphene electronics. *Nature Nanotechnol.* **5,** 722–726 (2010).
16. Novoselov, K. S. *et al.* Two-dimensional atomic crystals. *Proc. Natl Acad. Sci. USA* **102,** 10451–10453 (2005).
    This paper demonstrates that a number of 2D atomic crystals can be obtained in a free-standing state and used in various electronic devices.
17. Geim, A. K. Nobel lecture. Random walk to graphene. *Rev. Mod. Phys.* **83,** 851–862 (2011).
18. Novoselov, K. S. Nobel lecture. Graphene: materials in the flatland. *Rev. Mod. Phys.* **83,** 837–849 (2011).
19. Blake, P. *et al.* Graphene-based liquid crystal device. *Nano Lett.* **8,** 1704–1708 (2008).
20. Hernandez, Y. *et al.* High-yield production of graphene by liquid-phase exfoliation of graphite. *Nature Nanotechnol.* **3,** 563–568 (2008).
21. Coleman, J. N. *et al.* Two-dimensional nanosheets produced by liquid exfoliation of layered materials. *Science* **331,** 568–571 (2011).
22. Dreyer, D. R., Ruoff, R. S. & Bielawski, C. W. From conception to realization: an historical account of graphene and some perspectives for its future. *Angew. Chem. Int. Ed.* **49,** 9336–9344 (2010).
23. Schniepp, H. C. *et al.* Functionalized single graphene sheets derived from splitting graphite oxide. *J. Phys. Chem. B* **110,** 8535–8539 (2006).
24. Jiao, L. Y., Zhang, L., Wang, X. R., Diankov, G. & Dai, H. J. Narrow graphene nanoribbons from carbon nanotubes. *Nature* **458,** 877–880 (2009).
25. Kosynkin, D. V. *et al.* Longitudinal unzipping of carbon nanotubes to form graphene nanoribbons. *Nature* **458,** 872–876 (2009).
26. Segal, M. Selling graphene by the ton. *Nature Nanotechnol.* **4,** 612–614 (2009).
27. Bonaccorso, F., Sun, Z., Hasan, T. & Ferrari, A. C. Graphene photonics and optoelectronics. *Nature Photon.* **4,** 611–622 (2010).
28. Li, X. S. *et al.* Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **324,** 1312–1314 (2009).
    This paper indroduces CVD growth of graphene on copper, demonstrating the first large-area reproducible monolayer growth process.
29. Bae, S. *et al.* Roll-to-roll production of 30-inch graphene films for transparent electrodes. *Nature Nanotechnol.* **5,** 574–578 (2010).

30. Forbeaux, I., Themlin, J. M. & Debever, J. M. Heteroepitaxial graphite on 6H-SiC(0001): interface formation through conduction-band electronic structure. *Phys. Rev. B* **58,** 16396–16406 (1998).
31. Berger, C. *et al.* Ultrathin epitaxial graphite: 2D electron gas properties and a route toward graphene-based nanoelectronics. *J. Phys. Chem. B* **108,** 19912–19916 (2004).
32. Ohta, T., Bostwick, A., Seyller, T., Horn, K. & Rotenberg, E. Controlling the electronic structure of bilayer graphene. *Science* **313,** 951–954 (2006).
33. Virojanadara, C. *et al.* Homogeneous large-area graphene layer growth on 6H-SiC(0001). *Phys. Rev. B* **78,** 245403 (2008).
34. Lin, Y. M. *et al.* 100-GHz transistors from wafer-scale epitaxial graphene. *Science* **327,** 662 (2010).
This paper discusses the use of graphene epitaxially grown on SiC for high-frequency electronics.
35. Tzalenchuk, A. *et al.* Towards a quantum resistance standard based on epitaxial graphene. *Nature Nanotechnol.* **5,** 186–189 (2010).
36. Cai, J. M. *et al.* Atomically precise bottom-up fabrication of graphene nanoribbons. *Nature* **466,** 470–473 (2010).
37. Hackley, J., Ali, D., DiPasquale, J., Demaree, J. D. & Richardson, C. J. K. Graphitic carbon growth on Si(111) using solid source molecular beam epitaxy. *Appl. Phys. Lett.* **95,** 133114 (2009).
38. Dhar, S. *et al.* A new route to graphene layers by selective laser ablation. *AIP Adv.* **1,** 022109 (2011).
39. Han, T. H. *et al.* Extremely efficient flexible organic light-emitting diodes with modified graphene anode. *Nature Photon.* **6,** 105–110 (2012).
40. Liao, L. *et al.* High-speed graphene transistors with a self-aligned nanowire gate. *Nature* **467,** 305–308 (2010).
41. Liao, L. *et al.* Sub-100 nm channel length graphene transistors. *Nano Lett.* **10,** 3952–3956 (2010).
42. Han, S. J. *et al.* High-frequency graphene voltage amplifier. *Nano Lett.* **11,** 3690–3693 (2011).
43. Meric, I. *et al.* Channel length scaling in graphene field-effect transistors studied with pulsed current-voltage measurements. *Nano Lett.* **11,** 1093–1097 (2011).
44. Meric, I. *et al. High-Frequency Performance of Graphene Field Effect Transistors with Saturating IV-characteristics* 15–18 (IEEE Electron Devices Society, 2011).
45. Han, M. Y., Ozyilmaz, B., Zhang, Y. B. & Kim, P. Energy band-gap engineering of graphene nanoribbons. *Phys. Rev. Lett.* **98,** 206805 (2007).
46. Ponomarenko, L. A. *et al.* Chaotic Dirac billiard in graphene quantum dots. *Science* **320,** 356–358 (2008).
47. Stampfer, C. *et al.* Tunable graphene single electron transistor. *Nano Lett.* **8,** 2378–2383 (2008).
48. Oostinga, J. B., Heersche, H. B., Liu, X. L., Morpurgo, A. F. & Vandersypen, L. M. K. Gate-induced insulating state in bilayer graphene devices. *Nature Mater.* **7,** 151–157 (2008).
49. Kim, K., Choi, J. Y., Kim, T., Cho, S. H. & Chung, H. J. A role for graphene in silicon-based semiconductor devices. *Nature* **479,** 338–344 (2011).
50. Schwierz, F. Graphene transistors. *Nature Nanotechnol.* **5,** 487–496 (2010).
51. Britnell, L. *et al.* Field-effect tunneling transistor based on vertical graphene heterostructures. *Science* **335,** 947–950 (2012).
In this paper a new concept of vertical tunnelling transistors based on heterostructures assembled from 2D atomic crystals has been demonstrated.
52. Li, Z. Q. *et al.* Dirac charge dynamics in graphene by infrared spectroscopy. *Nature Phys.* **4,** 532–535 (2008).
53. Ishibashi, T. *et al.* InP/InGaAs uni-traveling-carrier photodiodes. *IEICE Trans. Electron. E* **83C,** 938–949 (2000).
54. Ishikawa, Y. & Wada, K. Near-infrared Ge photodiodes for Si photonics: operation frequency and an approach for the future. *IEEE Photon. J.* **2,** 306–320 (2010).
55. Xia, F. N., Mueller, T., Lin, Y. M., Valdes-Garcia, A. & Avouris, P. Ultrafast graphene photodetector. *Nature Nanotechnol.* **4,** 839–843 (2009).
This paper demonstrates the performance of planar graphene structures with built-in p–n junctions for ultrafast photodetection applications.
56. Meric, I. *et al.* Current saturation in zero-bandgap, topgated graphene field-effect transistors. *Nature Nanotechnol.* **3,** 654–659 (2008).
57. Xia, F. N. *et al.* Photocurrent imaging and efficient photon detection in a graphene transistor. *Nano Lett.* **9,** 1039–1044 (2009).
58. Mueller, T., Xia, F. N. A. & Avouris, P. Graphene photodetectors for high-speed optical communications. *Nature Photon.* **4,** 297–301 (2010).
59. Echtermeyer, T. J. *et al.* Strong plasmonic enhancement of photovoltage in graphene. *Nature Commun.* **2,** 458 (2011).
60. Reed, G. T., Mashanovich, G., Gardes, F. Y. & Thomson, D. J. Silicon optical modulators. *Nature Photon.* **4,** 518–526 (2010).
61. Liao, L. *et al.* 40 Gbit/s silicon optical modulator for highspeed applications. *Electron. Lett.* **43,** 1196–1197 (2007).
62. Li, G. L. *et al.* 25Gb/s 1V-driving CMOS ring modulator with integrated thermal tuning. *Opt. Express* **19,** 20435–20443 (2011).
63. Tang, Y. B. *et al.* 50 Gb/s hybrid silicon traveling-wave electroabsorption modulator. *Opt. Express* **19,** 5811–5816 (2011).
64. Wang, F. *et al.* Gate-variable optical transitions in graphene. *Science* **320,** 206–209 (2008).
65. Liu, M. *et al.* A graphene-based broadband optical modulator. *Nature* **474,** 64–67 (2011).
66. Sensale-Rodriguez, B. *et al.* Unique prospects for graphene-based terahertz modulators. *Appl. Phys. Lett.* **99,** 113104 (2011).
67. Liu, X., Du, D. & Mourou, G. Laser ablation and micromachining with ultrashort laser pulses. *IEEE J. Quantum Electron.* **33,** 1706–1716 (1997).
68. Drexler, W. *et al. In vivo* ultrahigh-resolution optical coherence tomography. *Opt. Lett.* **24,** 1221–1223 (1999).
69. Keller, U. *et al.* Semiconductor saturable absorber mirrors (SESAMs) for femtosecond to nanosecond pulse generation in solid-state lasers. *IEEE J. Quantum Electron.* **2,** 435–453 (1996).
70. Bao, Q. L. *et al.* Atomic-layer graphene as a saturable absorber for ultrafast pulsed lasers. *Adv. Funct. Mater.* **19,** 3077–3083 (2009).
71. Sun, Z. P. *et al.* Graphene mode-locked ultrafast laser. *ACS Nano* **4,** 803–810 (2010).
72. Zhang, H. *et al.* Graphene mode locked, wavelength-tunable, dissipative soliton fiber laser. *Appl. Phys. Lett.* **96,** 111112 (2010).
73. Xu, J. L. *et al.* Performance of large-area few-layer graphene saturable absorber in femtosecond bulk laser. *Appl. Phys. Lett.* **99,** 261107 (2011).
74. Tan, W. D. *et al.* Mode locking of ceramic Nd:yttrium aluminum garnet with graphene as a saturable absorber. *Appl. Phys. Lett.* **96,** 031106 (2010).
75. De Souza, E. A., Nuss, M. C., Knox, W. H. & Miller, D. A. B. Wavelength-division multiplexing with femtosecond pulses. *Opt. Lett.* **20,** 1166–1168 (1995).
76. Koch, B. R. *et al.* Mode locked and distributed feedback silicon evanescent lasers. *Laser Photon. Rev.* **3,** 355–369 (2009).
77. Rana, F. Graphene terahertz plasmon oscillators. *IEEE Trans. NanoTechnol.* **7,** 91–99 (2008).
78. Ramakrishnan, G., Chakkittakandy, R. & Planken, P. C. M. Terahertz generation from graphite. *Opt. Express* **17,** 16092–16099 (2009).
79. Prechtel, L. *et al.* Time-resolved ultrafast photocurrents and terahertz generation in freely suspended graphene. *Nature Commun.* **3,** 646 (2012).
80. Bao, Q. *et al.* Broadband graphene polarizer. *Nature Photon.* **5,** 411–415 (2011).
81. Bi, L. *et al.* On-chip optical isolation in monolithically integrated non-reciprocal optical resonators. *Nature Photon.* **5,** 758–762 (2011).
82. Crassee, I. *et al.* Giant Faraday rotation in single- and multilayer graphene. *Nature Phys.* **7,** 48–51 (2011).
83. Young, R. J., Kinloch, I. A., Gong, L. & Novoselov, K. S. The mechanics of graphene nanocomposites: a review. *Compos. Sci. Technol.* **72,** 1459–1476 (2012).
84. Wang, X., Zhi, L. J. & Mullen, K. Transparent, conductive graphene electrodes for dye-sensitized solar cells. *Nano Lett.* **8,** 323–327 (2008).
This described the first demonstration of the use of graphene (obtained via reduced graphene oxide method) as a transparent electrode in solar cells.
85. Li, S. S., Tu, K. H., Lin, C. C., Chen, C. W. & Chhowalla, M. Solution-processable graphene oxide as an efficient hole transport layer in polymer solar cells. *ACS Nano* **4,** 3169–3174 (2010).
86. Yang, S. B., Feng, X. L., Ivanovici, S. & Mullen, K. Fabrication of graphene-encapsulated oxide nanoparticles: towards high-performance anode materials for lithium storage. *Angew. Chem. Int. Edn* **49,** 8408–8411 (2010).
87. Yoo, E. *et al.* Large reversible Li storage of graphene nanosheet families for use in rechargeable lithium ion batteries. *Nano Lett.* **8,** 2277–2282 (2008).
88. Simon, P. & Gogotsi, Y. Materials for electrochemical capacitors. *Nature Mater.* **7,** 845–854 (2008).
89. Stoller, M. D., Park, S. J., Zhu, Y. W., An, J. H. & Ruoff, R. S. Graphene-based ultracapacitors. *Nano Lett.* **8,** 3498–3502 (2008).
This paper is the first demonstration of the use of graphene in a supercapacitor application.
90. Yoo, E. *et al.* Enhanced electrocatalytic activity of Pt subnanoclusters on graphene nanosheet surface. *Nano Lett.* **9,** 2255–2259 (2009).
91. Giesbers, A. J. M. *et al.* Quantum resistance metrology in graphene. *Appl. Phys. Lett.* **93,** 222109 (2008).
92. Nayak, T. R. *et al.* Graphene for controlled and accelerated osteogenic differentiation of human mesenchymal stem cells. *ACS Nano* **5,** 4670–4678 (2011).
93. Nair, R. R. *et al.* Graphene as a transparent conductive support for studying biological molecules by transmission electron microscopy. *Appl. Phys. Lett.* **97,** 153102 (2010).
94. Kuila, T. *et al.* Recent advances in graphene-based biosensors. *Biosens. Bioelectron.* **26,** 4637–4648 (2011).
95. Sanchez, V. C., Jachak, A., Hurt, R. H. & Kane, A. B. Biological interactions of graphene-family nanomaterials: an interdisciplinary review. *Chem. Res. Toxicol.* **25,** 15–34 (2012).
96. Yang, K. *et al.* Graphene in mice: ultrahigh *in vivo* tumor uptake and efficient photothermal therapy. *Nano Lett.* **10,** 3318–3323 (2010).
97. Nair, R. R., Wu, H. A., Jayaram, P. N., Grigorieva, I. V. & Geim, A. K. Unimpeded permeation of water through helium-leak-tight graphene-based membranes. *Science* **335,** 442–444 (2012).

Acknowledgements We are grateful to the graphene community for years of intensive research and discussions. In particular, A. Geim, F. Bonaccorso, I. Kinloch, R. J. Young, R. Dryfe, A. Tzalenchuk, D. Clarke, J. Kinaret and L. Eaves have commented on this paper. K.S.N. and V.I.F. acknowledge the EC Supporting Coordinated Action "Graphene-CA" Flagship Preparatory Action for financial support.

Author Contributions All authors contributed equally to the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence should be addressed to K.S.N. (kostya@manchester.ac.uk).

REVIEW

doi:10.1038/nature11320

Molecular machines governing exocytosis of synaptic vesicles

Reinhard Jahn[1] & Dirk Fasshauer[2]

Calcium-dependent exocytosis of synaptic vesicles mediates the release of neurotransmitters. Important proteins in this process have been identified such as the SNAREs, synaptotagmins, complexins, Munc18 and Munc13. Structural and functional studies have yielded a wealth of information about the physiological role of these proteins. However, it has been surprisingly difficult to arrive at a unified picture of the molecular sequence of events from vesicle docking to calcium-triggered membrane fusion. Using mainly a biochemical and biophysical perspective, we briefly survey the molecular mechanisms in an attempt to functionally integrate the key proteins into the emerging picture of the neuronal fusion machine.

Exocytosis and recycling of synaptic vesicles define how much transmitter is released from nerve terminals during incoming action potentials (Fig. 1). Under resting conditions, synaptic vesicles are stored in the cytoplasm of the nerve terminal, with some of them attached to specialized sites at the presynaptic plasma membrane termed active zones. Active zones are composed of unique multidomain proteins that provide a scaffold for vesicle docking and participate in activating the release apparatus, referred to as priming. Priming probably involves several reactions, including some requiring metabolic energy. Docked and primed vesicles (termed readily releasable pool) are ready to go, and some do so spontaneously, with the transmitter released by a single vesicle giving rise to a miniature postsynaptic potential. When an action potential arrives, voltage-gated calcium channels open, with the resulting calcium influx stimulating the rate of exocytosis more than 100,000 fold in a highly cooperative manner (for review see ref. 1).

During the past two decades, the key proteins mediating neuronal exocytosis have been identified. Many of them belong to structurally conserved protein families including the SNAREs, Rab proteins, Sec1/Munc18-like (SM) proteins, and a group of tethering proteins termed CATCHR (complex associated with tethering containing helical rods) proteins. Apparently, they form the core of an ancient intracellular fusion machine that diversified during evolution to adapt to the needs of specialized compartments. Neuronal exocytosis constitutes one of such adaptations, and specific regulatory proteins such as synaptotagmins and complexins evolved in the animal kingdom (for reviews see refs 2–8).

Despite such progress, there is still a gap in understanding between the functional properties of synaptic exocytosis and the molecular features of the key proteins. Modern electrophysiological[9,10] and imaging approaches[11–13] provided a wealth of information about the number of docked and primed vesicles, the exchange rates of vesicles between different pools, their release probabilities, their kinetics of exocytosis, and the dependence of exocytosis on calcium. Thus, detailed job descriptions for the underlying molecular machines are available. However, whereas genetic perturbations were instrumental in defining the basic functions of the key proteins, it often proved difficult to assign them to a specific step in the exocytotic pathway. For instance, Munc18, synaptotagmin and even the SNAREs were shown to function in docking as well as in priming and triggering. Conversely, specific steps such as docking are controlled by multiple proteins (see refs 7 and 11 for a more detailed discussion). It also often proved difficult to reconcile the physiological effects of the perturbations with the physicochemical properties of the proteins. Thus, the molecular mechanisms responsible for the attachment of synaptic vesicles to the active zone, for the activation of the release machinery, and for calcium triggering of exocytosis on a millisecond timescale are only slowly emerging.

SNARE proteins, the engine of membrane fusion

The synaptic proteins synaptobrevin (also referred to as VAMP), syntaxin 1 and SNAP-25 belong to the SNARE protein family. Their

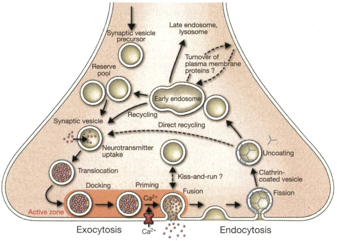


Figure 1 | Trafficking pathways in the nerve terminal. Synaptic vesicles are filled with neurotransmitter and stored in the cytoplasm. Active vesicles are translocated to release sites in the active zone where they dock. Priming involves all steps required to acquire release readiness of the exocytotic complex. Although usually assumed to occur after docking, priming and even triggering may precede docking during sustained activity, resulting in immediate fusion of an arriving vesicle. After exocytosis, the vesicle proteins probably remain clustered and are then retrieved by endocytosis. Despite some lingering controversies, consensus is emerging that retrieval is generally mediated by clathrin-mediated endocytosis. After clathrin uncoating, synaptic vesicles are regenerated within the nerve terminal, probably involving passage through an endosomal intermediate. Actively recycling vesicles are in slow exchange with the reserve pool. See text for more details.

---

[1] Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany. [2] Department of Basic Neuroscience, Faculty of Biology and Medicine, University of Lausanne, 1005 Lausanne, Switzerland.

defining feature is an extended coiled-coil stretch, which is referred to as a SNARE motif and falls into four subtypes, referred to as Qa, Qb, Qc and R-SNARE motif (for example ref. 14). In syntaxin, synaptobrevin and in most other SNAREs the SNARE motifs are connected by a short linker to a carboxy-terminal transmembrane region (TMR). SNAP-25 deviates from this general structure: here two SNARE motifs (Qb and Qc) are connected by a linker that is palmitoylated, whereas a TMR is lacking. Whereas synaptobrevin and SNAP-25 do not carry any other domains, syntaxin possesses an amino-terminal domain consisting of an antiparallel three-helix bundle, termed the Habc domain[15,16], connected to the SNARE motif by a flexible linker. Positioned N-terminally to the Habc domain is a short stretch that ends in the so-called N-peptide (see Fig. 2).

SNAREs undergo a regulated assembly–disassembly cycle that is energized by the AAA+-ATPase NSF. Synaptobrevin is a synaptic vesicle protein, whereas syntaxin 1 and SNAP-25 are localized in the presynaptic plasma membrane. On contact, the SNAREs associate in *trans* at the N-terminal ends of the SNARE motifs. A tight bundle of four parallel α-helices is formed, each contributed by a different SNARE motif[17,18], which progresses towards the C-terminal membrane anchors ('zippering'), thus pulling the membranes tightly together[19]. Assembly is associated with a huge release of energy that is used to initiate membrane fusion[20,21]. After fusion, the ternary SNARE complex resides in the plasma membrane in the low-energy *cis* configuration and is disassembled by NSF in conjunction with its SNAP cofactor. Next, synaptobrevin is endocytosed and recycled, thus being able to participate in another round of exocytosis (for reviews see refs 2–8).

Despite the elegant simplicity and experimental support[22] of the zippering model, SNARE assembly proved to be an unexpectedly complex reaction, and there is still a lot to learn. *In vitro*, isolated SNARE motifs are unfolded but assemble into diverse homo- and hetero-oligomers that all are at least partially helical (reviewed in refs 4 and 8). For instance, SNAP-25 can bind sequentially two syntaxin molecules, thus blocking the binding site of synaptobrevin[23]. Furthermore, syntaxin rapidly switches between an active open conformation and an inactive closed conformation in which the Habc domain folds against the N-terminal part of the SNARE motif[24,25]. Such conformational dynamics and kinetic trapping of off-pathway intermediates explains why *in vitro* assembly of the ternary complex, although highly exergonic, lasts hours, far too slow to mediate fast exocytosis. On the other hand, if a complex of SNAP-25 and syntaxin with a free N-terminal binding site for synaptobrevin is stabilized, SNARE assembly is accelerated by orders of magnitude[26]. The central problem is to delineate precisely the assembly pathway and to understand how the SNARE molecules are channelled along this pathway by regulatory proteins to execute fusion efficiently. Four proteins, each representing a small protein family, have emerged as such key regulators: Munc18 and Munc13 that prepare the SNARE engine for assembly, and synaptotagmin and complexin that govern calcium-dependent triggering.

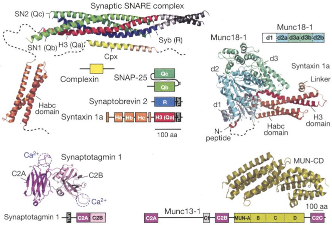


Figure 2 | Schematic depictions of domain structures and crystal structures of core proteins of the neuronal fusion machine. The dashed lines between the N-peptide (N) and Habc domain represent flexible regions in syntaxin. For synaptotagmin the two Ca^{2+} binding sites are indicated. Note that the domain structure of the large multi-domain protein Munc13 is shown five times smaller than those of the other domain structures. A high-resolution structure was obtained for the C-terminal half of the MUN domain. See text for details. The data for structures are from: Munc13-1 (C and D subdomains)[48], Synaptotagmin 1 (C2A and C2B domain)[117], Munc18-1 (blue-green, in complex with syntaxin (red))[32], Habc domain[16], SNARE complex[18], complexin[63]. aa, amino acid.

Priming the SNARE engine

UNC-18 and UNC-13, the *Caenorhabditis elegans* orthologues of Munc18 and Munc13, respectively, were originally identified by S. Brenner in his classical screen uncovering genes involved in movement[27]. Deletion of either Munc18 (ref. 28) or Munc13 (ref. 29) and their respective orthologues[30,31] completely inhibits neuronal exocytosis. Munc18 belongs to the conserved family of SM proteins. It possesses an arch-shaped architecture with a central cavity for high-affinity binding to syntaxin-1 (refs 32, 33). By contrast, the large Munc13s belong to the CATCHR protein family[34]. Munc13 also binds to syntaxin-1 but only with moderate affinity[35,36]. Both proteins are involved in setting up the SNAREs for assembly and perhaps in guiding them through the initial part of the assembly pathway, but it is still not understood how exactly they operate, how many copies are required to carry out the reaction, and how the extraordinary phenotypes of the knockouts can be mechanistically explained.

Munc18

For many years, the molecular mechanism of Munc18 has been shrouded by a paradox because it locks syntaxin-1 in a closed conformation[24,33] (Fig. 2), in which syntaxin cannot enter SNARE complexes. Such inhibition is difficult to reconcile with the complete loss of exocytosis in deletion mutants, which suggests exactly the opposite, namely that SNARE zippering is absolutely dependent on Munc18. Indeed, Munc18 seems to be an oddity because other SM proteins, despite high structural similarity, bind instead tightly to the N-peptide of their cognate syntaxins, involving a binding site on the surface of the SM protein. This binding mode would enable these syntaxins to remain open, with SNARE assembly not being inhibited, whereas syntaxin-1 would need to be opened in the case of Munc18. To reconcile these discrepancies, it was proposed that binding of SM proteins to syntaxins, whether via the N-peptide or the Habc domain, merely serves to recruit the SM protein to the prospective fusion site. The SM proteins are then handed over to the SNARE motifs where they promote nucleation and/or zippering (for reviews see refs 3 and 37).

Recently, it has been recognized that SM proteins, including Munc18, generally bind to their respective syntaxins using both of the spatially distinct binding sites, but with different relative affinities[32,38,39]. In fact, the two binding sites seem to act together in controlling SNARE complex formation[32]. This sheds new light on the paradox, as full 'opening' of syntaxin may not be required for gating entry into SNARE complexes. In support of this view, a syntaxin mutant originally thought to be constitutively open (LE mutant)[24] is now known to bind Munc18 via both sites in an at least partially closed conformation, but without inhibiting formation of SNARE complexes[32,36]. Indeed, when expressed as the only syntaxin 1 variant, the LE mutant results in enhanced spontaneous exocytosis, supporting that under resting conditions it is more reactive with respect to SNARE binding[40].

Thus it seems that binding of Munc18-1 to both the closed conformation and to the N-peptide of syntaxin 1a is an integral part of the pathway during which Munc18 guides syntaxin towards productive SNARE complex formation. Perhaps Munc18 first keeps syntaxin closed and inactive, thus preventing premature SNARE assembly, but allows for synchronization of a subsequent (calcium-dependent?) activation step (see for example ref. 3).

Despite such progress, it is still unclear why Munc18 is essential for efficient SNARE nucleation. Reconstitution experiments involving liposomes suggest that Munc18 participates in selecting the correct R-SNARE helix and thus guides nucleation of the ternary complex (for example refs 41 and 42). It is unclear whether it then remains associated with the ternary complex[43] as also suggested for other SM proteins (for example ref. 37) or dissociates upon zippering[44]. In the latter case Munc18 might be interacting with a *trans*-SNARE complex only during its initiation, whereas progression of zippering would cause syntaxin to fully open, thus driving off Munc18.

Munc13

Munc13s are modular proteins sharing a conserved C-terminal region containing a phorbol-ester-binding C1 domain and two calcium-binding C2 domains that flank a larger, so-called MUN domain (Fig. 2, reviewed in ref. 45). Expression of the MUN domain alone partially rescues the total arrest of exocytosis in neurons lacking Munc13s[46], identifying it as a key functional element of the protein. MUN domains are shared with the proteins BAP3 and CAPS (Unc31) and with other proteins in most eukaryotes[47]. The MUN domain is structurally strikingly similar to other CATCHR family members[48] that work in various trafficking steps. These proteins form elongated arrays of stacked α-helical bundles with flexible hinge regions, which tether transport vesicles to the site of fusion. It is conceivable that the conserved MUN domain serves as binding platforms that arrange the core fusion machinery, whereas the C1 and C2 domains mediate fine-tuning of its membrane recruitment, a feature ideally suited to Ca^{2+}-regulated secretion.

CATCHR complexes are also thought to enable SNARE assembly, although their interplay with SNARE proteins seems to vary. For instance, Munc13 may participate directly by unlocking syntaxin from the grip of Munc18, because in *C. elegans*, the LE mutant of syntaxin (that is not inhibited by Munc18 binding) partially rescues neurotransmitter release in the absence of UNC-13 (refs 49 and 50). Furthermore, recent experiments have shown that the isolated MUN domain accelerates the transition of syntaxin-1 from the Munc18-1 complex to the SNARE complex[36]. It should be kept in mind, however, that the LE mutant also partially rescues the block of exocytosis caused by deletion of RIM (also known as UNC-10) in *C. elegans*[51]. RIM serves as central organizer of the active zone. It forms a tripartite complex with the N-terminal C2A domain of Munc13 and the small vesicular GTPases RAB3 and RAB27, thus orchestrating the attachment site of synaptic vesicles (reviewed in ref. 2).

Ca^{2+}-dependent triggering starts the SNARE engine

In contrast to the basic fusion reaction that is carried out by conserved proteins traced back to an ancient eukaryotic machine, the unique features of calcium-triggered exocytosis are primarily encoded in specialized proteins. Of these, synaptotagmins I, II and IX constitute the dominant calcium sensors whose deletion results in a complete loss of fast, calcium-triggered exocytosis (reviewed in refs 52–54). However, asynchronous (that is, slower) calcium-dependent release persists, showing that other calcium-binding proteins are involved, with candidates including other synaptotagmin isoforms or related proteins such as Doc2 (refs 55–57). Furthermore, complexins I and II are involved in triggering: deletion of complexins strongly reduces calcium-evoked exocytosis, whereas both stimulatory and inhibitory effects were observed on spontaneous release (for example see refs 58 and 59).

Synaptotagmins

The neuronal synaptotagmins are anchored to synaptic vesicles by a single TMR. Characteristic features of the synaptotagmins are two C2 domains, called C2A and C2B, that are connected to the membrane as well as to each other by flexible linkers. C2 domains are rigid, oval-shaped β-sandwiches that possess a cluster of calcium-binding loops, serving as partial coordination site for two (C2A) or three (C2B) calcium ions. In the presence of calcium, the C2 domains bind to membranes containing acidic phospholipids that complete the calcium coordination sites. In addition, the C2B domain contains a spatially separated basic patch that steers the domain to membranes enriched in phosphatidylinositol (4,5) bisphosphate ($PI(4,5)P_2$). Membrane binding is primarily electrostatic and rapidly reversed by chelating calcium or increasing the ionic strength. Furthermore, the synaptotagmin C2 domains bind to syntaxin alone or syntaxin-containing SNARE complexes (for example see refs 60 and 61). Although binding occurs in the absence of calcium, it appears to be influenced by calcium (reviewed in refs 54 and 62).

Complexins

Complexins are small cytoplasmic proteins that bind via a central helix to a groove on the surface of the SNARE complex, which is formed by the helices of syntaxin and synaptobrevin[63,64] (Fig. 2). Because SNARE binding is required for their physiological action, complexins can only exert their function once SNAREs are at least partially assembled, placing them into the reaction sequence after zippering is initiated. Intriguingly, the central helix is not sufficient for complexin function. Rather, the N-terminal end pointing towards the membrane is needed for facilitation of fusion, whereas the regions flanking the central helix seem to have an inhibitory role. To accommodate the presumed dual stimulatory and inhibitory role of complexin, two alternative molecular mechanisms are discussed[3,9,65–67]. First, binding to the surface of the SNARE complex may promote initiation and progression of zippering, for example, by stabilizing partially zippered SNARE complexes and sensitizing them to activation by synaptotagmin ('super-priming') (for example see ref. 58). Second, complexin acts as a clamp that blocks progression of SNARE-zippering, presumably by competing directly with synaptobrevin binding in the C-terminal part of the SNARE complex (for example see refs 59 and 68). The clamp is released upon calcium triggering, probably by synaptotagmin (see below) because complexins do not bind calcium.

Two models explain the action of calcium

Despite many years of research, it is still controversial as to how calcium influx brings about the extraordinary and highly cooperative acceleration of exocytosis. To some extent this is owing to the fact that the molecular status of a docked and primed vesicle, ready to respond to calcium by exocytosis in less than a millisecond, is not known with certainty.

Most authors seem to agree that SNAREs are already partially zippered in this state, with full zippering being prevented either by an energy barrier in the fusion pathway that the SNAREs alone cannot overcome (for example, electrostatic repulsion, transition towards a stalk intermediate, see below), and/or by an interfering protein, with prime candidates being complexins and/or synaptotagmins (Fig. 3, pathway I). During this state, Munc18, and perhaps also Munc13, may still be bound to the complex. It is debated whether such a complex is strained, that is, storing energy that is released during fusion, or whether it is relaxed, with the linkers connecting the zippered part of the complex to the membrane being flexible.

Calcium binding to synaptotagmin would trigger fusion either by activating (disinhibiting) the SNAREs or by lowering the activation energy barrier in the fusion pathway through membrane interactions. Accordingly, synaptotagmin may act by (1) disengaging from the SNAREs, thus relieving the block (fusion clamp model)[69], (2) binding to the SNAREs, thus displacing the inhibitory complexin and/or promoting zippering[59], (3) binding to the membrane directly adjacent to the partially complexed SNAREs, thus destabilizing the bilayer at the fusion site[70–72], (4) increasing curvature stress by displacing lipids in the monolayer of the plasma membrane facing the vesicle[73,74], and (5) cross-linking the vesicle and the plasma membrane, thus accelerating fusion by charge compensation owing to the positive electrostatic potential of the C2 domains[75].

A wealth of evidence is invoked in support of a partially zippered and arrested SNARE complex; for instance, differential effects of SNARE mutations on fusion kinetics that affect nucleation and zippering,

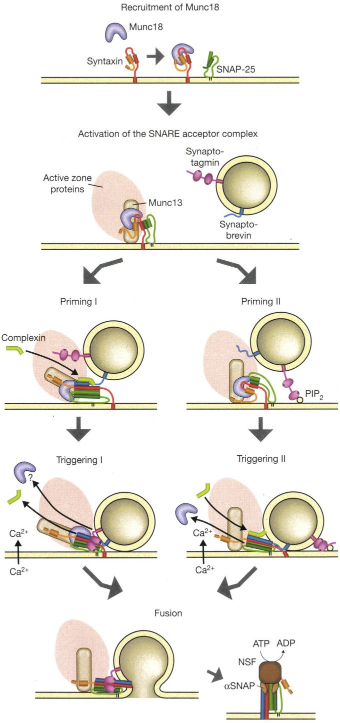


Figure 3 | Alternative models describing the steps between priming and fusion. Priming I involves arrest of a partially zippered SNARE complex, here shown with bound Munc18, Munc13 and synaptotagmin. Calcium influx triggers binding of synaptotagmin to the SNARE complex and to the plasma membrane (involving PI(4,5)P_2, not shown here), associated with displacement of complexin and (possibly) Munc18 and/or Munc13. Priming II involves arrest after positioning of the vesicle with the aid of active zone components and (possibly) contact of synaptotagmin with PI(4,5)P_2 in the plasma membrane, but no contact between the SNAREs. $\mathrm{Ca^{2+}}$-triggering pulls the vesicle closer via synaptotagmin-mediated cross-linking, resulting in SNARE assembly, associated with full opening of syntaxin and displacement of Munc18, and binding of complexin. See text for details.

respectively (see for example refs 76–78). The model also allows for an integration of complexin into the fusion mechanism that needs at least partial SNARE assembly before it can bind and exert its action. Furthermore, the model intuitively explains the fast fusion kinetics because only minor conformational rearrangements are required upon $\mathrm{Ca^{2+}}$-triggering, with all proteins already being correctly positioned for the final step.

On the other hand, there are problems with this model, which in our opinion have not been sufficiently appreciated. Experimentally, *trans*-complexes are difficult to capture. Similarly, despite hints, for example, from single-molecule experiments[4,79], an effect of synaptotagmin and/or complexin on the rate of SNARE assembly has remained elusive. Most importantly, the mechanisms proposed for arresting SNARE zippering somewhere in the middle are difficult to reconcile with the fact that SNARE assembly proceeds along a steep downhill energy gradient. For instance, a C-terminal fragment of synaptobrevin forms *in vitro* a stable complex with SNAP-25 and syntaxin, thus blocking, like a brake shoe, the C-terminal portion from assembling as envisioned in the partially zippered model. However, full-length synaptobrevin is able to rapidly displace this fragment[26] and, despite this additional energy barrier, to promote fusion *in vitro* even if only one of such partially inhibited SNARE complexes is involved[80]. None of the proposed factors (including complexin) binds with an affinity even remotely comparable to that of the synaptobrevin fragment, questioning their ability to influence the strongly exergonic zippering reaction. Furthermore, we consider it unlikely that the control of the neuronal SNAREs is exerted by tinkering with its structurally highly conserved engine core, that is, the helical bundle whose major features (structure, stability, folding–unfolding hysteresis) are remarkably similar between SNAREs in regulated and non-regulated trafficking steps[8].

Similarly, a role of synaptotagmin in destabilizing membranes or inducing curvature stress is difficult to reconcile with $\mathrm{Ca^{2+}}$-dependent membrane binding being primarily electrostatic and reversible. Indeed, vesicle deformation *in vitro* requires saturation with C2 domain-containing synaptotagmin fragments[73,74], whereas only few phospholipid molecules are expected to be displaced by binding of single C2 domains, hardly sufficient to create even local tension. Generally, agents increasing positive spontaneous curvature of the proximal monolayers inhibit rather than enhance fusion[81], that is, exactly the opposite of what synaptotagmin is doing, requiring elaborate models of highly organized 'bulges' to explain promotion of fusion intermediates[82].

Recently, an alternative scenario for the docked and primed state has been envisaged that, despite being far from proven, we consider as an interesting alternative, as it overcomes several of the problems outlined above[70,83] (Fig. 3, pathway II, see also ref. 40). Supported by recent electron tomography data showing that docked vesicles appear to be a few nanometres away from the plasma membrane[11,12], it assumes that SNAREs do not connect *in trans* before the arrival of the calcium signal. Rather, the interaction of the vesicle with the active zone components (most notably Munc13 and RIM) would precisely position the vesicle on top of a patch of plasma membrane containing activated SNARE acceptor complexes, probably complexed with Munc18. In this state, vesicle-bound synaptotagmin may be already in contact with the plasma membrane, either by (calcium-independent) binding to the SNAREs or by loosely binding to PI(4,5)P_2 patches colocalizing with syntaxin clusters. Calcium influx would trigger membrane binding and cross-linking of the vesicle and plasma membrane, thus nudging them a bit closer[75,83,84], sufficient to allow for rapid binding of synaptobrevin to the acceptor complex[78]. Once nucleation is triggered, the SNAREs quickly progress through zippering and fusion.

This model places the entire control of the neuronal fusion machine upstream of SNARE nucleation, which has important consequences for our understanding of the partial reactions. Most importantly, it changes the view of SNARE function. Accordingly, SNAREs act as 'single shot' devices that, once nucleation is triggered, are unstoppable and flash through assembly to bring about fusion. 'Misfiring' of SNAREs (assembly without fusion) probably only occurs rarely, if at all, but is likely to increase in mutants affecting zippering[85]. Also, it is possible that nucleation triggers the displacement of Munc18 and other factors (such as Munc13), thus allowing the SNAREs to carry out the work unhindered by bulky bound proteins. Such a simple and highly efficient mode of operation may explain why the SNARE engine was so successful in evolution. The proposed function for synaptotagmin is in line with the function of C2 domains in other proteins such as protein kinase C—they operate as electrostatic switches[62] mediating calcium-dependent rapid and reversible membrane binding.

The model also elegantly explains why solutions of high osmolarity (usually sucrose) trigger calcium-independent exocytosis of the readily releasable pool[86]: the resulting water efflux creates negative pressure that draws docked vesicles closer to the plasma membrane, triggering SNARE firing. Furthermore, any destabilization of the overall architecture of the docking site, which increases Brownian fluctuations of the vesicle, would cause occasional spontaneous firing of SNAREs, which may explain changes in spontaneous release rates upon deletion or overexpression of some proteins (for a detailed discussion see for example refs 9 and 57).

Finally, the model provides for a fresh look at the molecular basis of the high cooperativity of calcium-triggered fusion. At non-saturating calcium concentrations, synaptotagmin binding may be less tight or transient, perhaps undergoing rapid and repetitive 'on–off' cycles, resulting in vesicle jittering. Accordingly, the probability for SNARE nucleation/firing would be reduced. Such a scenario may also explain the function of complexin, which is otherwise more difficult to integrate into this model. Complexin may increase the frequency of successful nucleation events by stabilizing correctly oriented syntaxin–synaptobrevin alignments. This hypothesis is in line with the ability of complexin to bind to SNARE complexes with fast, diffusion-limited kinetics[87].

Fusion—interplay between proteins and lipids

In the final step of exocytosis the vesicle membrane fuses with the plasma membrane. The merger of two bilayers involves non-bilayer intermediates at the contact site that ultimately develop into the opening of an aqueous channel, termed a fusion pore. During fusion the hydrophobic barrier separating the cytoplasm from both the vesicle content and the extracellular space must remain intact.

Key issues concerning the molecular rearrangements of proteins and membrane lipids along the fusion pathway are unresolved. Popular models requiring an oligomeric ring of SNARE complexes surrounding the prospective fusion pore as intermediate cannot be maintained in view of the fact that only one to two (or three) SNARE complexes are sufficient for fusion both *in vitro* and *in vivo*[80,88–90]. Intriguingly, *in vitro* fusion can be mediated by *trans* assembly of artificially engineered molecules mimicking SNARE-zippering (even DNA) as long as they possess membrane anchors (for example see refs 91–94). Such a lack of structural specificity in catalysis is indeed a hallmark of membrane fusion, and it is likely that considerable structural variety is tolerated along the fusion pathway. This helps to explain why unrelated classes of fusion proteins evolved in parallel to the SNAREs, such as those fusing cells[95], viruses[96], mitochondria[97] or the endoplasmic reticulum[98].

The stalk hypothesis, first developed 30 years ago[99], describes membrane fusion as an ordered sequence of steps initiated by an hourglass-shaped intermediate (the fusion stalk), followed by a hemifusion diaphragm and subsequent rupture, resulting in the formation of a fusion pore (Fig. 4). Indeed, stalk-like intermediates can be induced as a separate phase under mild conditions[100–102]. However, the energy landscape as well as the intermediate molecular structures along the fusion pathway is unclear.

Originally, the energy profile was modelled on the basis of the elastic properties of membranes, with the curvature stress of the intermediate model structures defining transition-state energies. However, these energies were unrealistically high, and molecular parameters were invoked to lower the energies (for review see ref. 103). More recently, coarse-grain or even atomistic simulations of fusion have provided detailed scenarios for intermediate structures (Fig. 4), with consequences for the energy landscape. For instance, it has been suggested that 'splaying' of phospholipid tails may form the first hydrophobic connection between the membranes[104] from where stalk formation proceeds downhill an energy gradient. Furthermore, the enhanced fusogenicity of curved membranes can also be explained by the hydrophobic effect: owing to the increased spacing of the hydrophilic head groups the membrane surface is more hydrophobic. Lipid splaying requires the membranes to be at a critical distance of less than 1 nm (Fig. 4, see ref. 105 for a more detailed discussion).

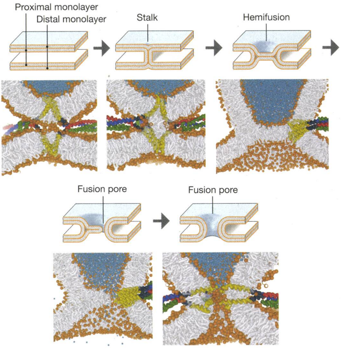


Figure 4 | Transition states during membrane fusion. Intermediates of the fusion pathway. The top drawings represent intermediate states of the membrane along the pathway as predicted by the elastic theory. Below, snapshots of intermediate states of a simulation of SNARE-mediated fusion are shown, which, although roughly corresponding to the elastic model, differ in detail and in their energy predictions (adapted from ref. 109, courtesy of J. Risselada and H. Grubmüller).

These considerations have important consequences for the mechanism of SNARE-catalysed fusion. Certainly, zippering of the four-helix bundle brings the membranes in close proximity, but the question is how the SNAREs promote stalk formation and subsequent intermediate structures. If the main energy barrier is contributed by curvature stress[106], stiffness of the linkers connecting SNARE motifs and TMRs is essential for transmitting stress to the membranes. Indeed, mutagenesis of the linkers generally reduces fusion efficacy (see for example ref. 107), and at least syntaxin seems to have a stiff linker as a monomer[108]. On the other hand, if close proximity, water removal, increase of local hydrophobicity and lipid splaying form the main energy barrier, bending stiffness of the SNARE linkers may not be as relevant, a view suggested by recent simulation studies[109]. Instead, the pulling force exerted during zippering may drag the TMRs along with some phospholipids slightly out of the membrane, thus initiating phospholipid splaying once the critical distance has been reached.

What are the next steps? Transient hemifusion intermediates (experimentally defined, for example, by lipid mixing in the absence of content mixing) are observed upon SNARE-mediated fusion of liposomes, suggesting hemifusion as a metastable intermediate (for example see refs 110–114). However, it is experimentally difficult to differentiate between stalk and hemifusion intermediates. Hemifusion constitutes the lateral expansion of a stalk, leading to the formation of a hemifusion diaphragm (Fig. 4). It remains to be seen whether such diaphragms represent intermediates along the fusion pathway or whether they are dead-ends as previously suggested for viral fusion proteins (see ref. 115

for a review). In any case, the job of the SNAREs is not finished before the initial opening of the fusion pore, with interactions between the linkers as well as the TMRs probably being involved[116].

Conclusion

The molecular basis of synaptic exocytosis has fascinated scientists for decades. Since the initial discovery of quantal release in the 1950s by Katz and colleagues, and the elucidation of the synaptic vesicle-recycling pathway by Heuser and Ceccarelli in the 1970s, we have come a long way in deciphering the steps of the vesicle cycle at an increasingly detailed level. Although we have focused here on only a few key components, the vesicle cycle is governed by hundreds of proteins, and there are still new proteins being put on the map. We are only beginning to understand the rules by which individual protein–protein interactions work together in supramolecular machines to yield the synaptic vesicle cycle that reliably operates millions of times. These machines assemble on demand and disassemble when the task is completed. They are highly robust, tolerate varying stoichiometries, flexible compositions and other disturbances, and are controlled by an array of regulators such as protein kinases and phosphatases. Advances in technologies such as super-resolution microscopy, single-molecule measurements, fluorescent reporters and cryo-electron tomography are all contributing to closing the gap between our understanding of partial reactions *in vitro* and the fascinating efficiency of the vesicle cycle in intact synapses.

1. Südhof, T. C. The synaptic vesicle cycle. *Annu. Rev. Neurosci.* **27,** 509–547 (2004).
2. Südhof, T. C. & Rizo, J. Synaptic vesicle exocytosis. *Cold Spring Harb. Perspect. Biol.* **3,** a005637 (2011).
3. Südhof, T. C. & Rothman, J. E. Membrane fusion: grappling with SNARE and SM proteins. *Science* **323,** 474–477 (2009).
4. Brunger, A. T., Weninger, K., Bowen, M. & Chu, S. Single-molecule studies of the neuronal SNARE fusion machinery. *Annu. Rev. Biochem.* **78,** 903–928 (2009).
5. Wickner, W. & Schekman, R. Membrane fusion. *Nature Struct. Mol. Biol.* **15,** 658–664 (2008).
6. Rizo, J. & Rosenmund, C. Synaptic vesicle fusion. *Nature Struct. Mol. Biol.* **15,** 665–674 (2008).
7. Verhage, M. & Sørensen, J. B. Vesicle docking in regulated exocytosis. *Traffic* **9,** 1414–1424 (2008).
8. Jahn, R. & Scheller, R. H. SNAREs–engines for membrane fusion. *Nature Rev. Mol. Cell Biol.* **7,** 631–643 (2006).
9. Sørensen, J. B. Conflicting views on the membrane fusion machinery and the fusion pore. *Annu. Rev. Cell Dev. Biol.* **25,** 513–537 (2009).
10. Neher, E. & Sakaba, T. Multiple roles of calcium ions in the regulation of neurotransmitter release. *Neuron* **59,** 861–872 (2008).
11. Siksou, L., Triller, A. & Marty, S. Ultrastructural organization of presynaptic terminals. *Curr. Opin. Neurobiol.* **21,** 261–268 (2011).
12. Fernáandez-Busnadiego, R. *et al.* Insights into the molecular organization of the neuron by cryo-electron tomography. *J. Electron Microsc. (Tokyo)* **60** (suppl. 1), S137–S148 (2011).
13. Sigrist, S. J. & Sabatini, B. L. Optical super-resolution microscopy in neurobiology. *Curr. Opin. Neurobiol.* **22,** 86–93 (2012).
14. Kloepper, T. H., Kienle, C. N. & Fasshauer, D. An elaborate classification of SNARE proteins sheds light on the conservation of the eukaryotic endomembrane system. *Mol. Biol. Cell* **18,** 3463–3471 (2007).
15. Fernandez, I. *et al.* Three-dimensional structure of an evolutionarily conserved N-terminal domain of syntaxin 1A. *Cell* **94,** 841–849 (1998).
16. Lerman, J. C., Robblee, J., Fairman, R. & Hughson, F. M. Structural analysis of the neuronal SNARE protein syntaxin-1A. *Biochemistry* **39,** 8470–8479 (2000).
17. Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395,** 347–353 (1998).
18. Stein, A., Weber, G., Wahl, M. C. & Jahn, R. Helical extension of the neuronal SNARE complex into the membrane. *Nature* **460,** 525–528 (2009).
This paper describes the X-ray structure of the synaptic SNARE complex with transmembrane regions.
19. Hanson, P. I., Heuser, J. E. & Jahn, R. Neurotransmitter release — four years of SNARE complexes. *Curr. Opin. Neurobiol.* **7,** 310–315 (1997).
20. Li, F. *et al.* Energetics and dynamics of SNAREpin folding across lipid bilayers. *Nature Struct. Mol. Biol.* **14,** 890–896 (2007).
21. Wiederhold, K. & Fasshauer, D. Is assembly of the SNARE complex enough to fuel membrane fusion? *J. Biol. Chem.* **284,** 13143–13152 (2009).
22. Weber, T. *et al.* SNAREpins: minimal machinery for membrane fusion. *Cell* **92,** 759–772 (1998).
23. Fasshauer, D. & Margittai, M. A Transient N-terminal interaction of SNAP-25 and syntaxin nucleates SNARE assembly. *J. Biol. Chem.* **279,** 7613–7621 (2004).
24. Dulubova, I. *et al.* A conformational switch in syntaxin during exocytosis: role of munc18. *EMBO J.* **18,** 4372–4382 (1999).
25. Margittai, M. *et al.* Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc. Natl Acad. Sci. USA* **100,** 15516–15521 (2003).
26. Pobbati, A. V., Stein, A. & Fasshauer, D. N- to C-terminal SNARE complex assembly promotes rapid membrane fusion. *Science* **313,** 673–676 (2006).
27. Brenner, S. The genetics of *Caenorhabditis elegans. Genetics* **77,** 71–94 (1974).
28. Verhage, M. *et al.* Synaptic assembly of the brain in the absence of neurotransmitter secretion. *Science* **287,** 864–869 (2000).
29. Varoqueaux, F. *et al.* Total arrest of spontaneous and evoked synaptic transmission but normal synaptogenesis in the absence of Munc13-mediated vesicle priming. *Proc. Natl Acad. Sci. USA* **99,** 9037–9042 (2002).
30. Richmond, J. E., Davis, W. S. & Jorgensen, E. M. UNC-13 is required for synaptic vesicle fusion in *C. elegans. Nature Neurosci.* **2,** 959–964 (1999).
31. Aravamudan, B., Fergestad, T., Davis, W. S., Rodesch, C. K. & Broadie, K. *Drosophila* UNC-13 is essential for synaptic transmission. *Nature Neurosci.* **2,** 965–971 (1999).
32. Burkhardt, P., Hattendorf, D. A., Weis, W. I. & Fasshauer, D. Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide. *EMBO J.* **27,** 923–933 (2008).
This paper reports that Munc18-1 interacts with two spatially separated binding sites of syntaxin-1a.
33. Misura, K. M., Scheller, R. H. & Weis, W. I. Three-dimensional structure of the neuronal-Sec1–syntaxin 1a complex. *Nature* **404,** 355–362 (2000).
34. Yu, I. M. & Hughson, F. M. Tethering factors as organizers of intracellular vesicular traffic. *Annu. Rev. Cell Dev. Biol.* **26,** 137–156 (2010).
35. Betz, A., Okamoto, M., Benseler, F. & Brose, N. Direct interaction of the rat *unc-13* homologue Munc13–1 with the N terminus of syntaxin. *J. Biol. Chem.* **272,** 2520–2526 (1997).
36. Ma, C., Li, W., Xu, Y. & Rizo, J. Munc13 mediates the transition from the closed syntaxin-Munc18 complex to the SNARE complex. *Nature Struct. Mol. Biol.* **18,** 542–549 (2011).
37. Toonen, R. F. & Verhage, M. Munc18–1 in secretion: lonely Munc joins SNARE team and takes control. *Trends Neurosci.* **30,** 564–572 (2007).
38. Furgason, M. L. *et al.* The N-terminal peptide of the syntaxin Tlg2p modulates binding of its closed conformation to Vps45p. *Proc. Natl Acad. Sci. USA* **106,** 14303–14308 (2009).
39. Khvotchev, M. *et al.* Dual modes of Munc18-1/SNARE interactions are coupled by functionally critical binding to syntaxin-1 N terminus. *J. Neurosci.* **27,** 12147–12155 (2007).
40. Gerber, S. H. *et al.* Conformational switch of syntaxin-1 controls synaptic vesicle fusion. *Science* **321,** 1507–1510 (2008).
This paper and ref. 49 describe the complex phenotype of the LE mutant of syntaxin on docking and fusion of synaptic vesicles.
41. Rathore, S. S. *et al.* Syntaxin N-terminal peptide motif is an initiation factor for the assembly of the SNARE-Sec1/Munc18 membrane fusion complex. *Proc. Natl Acad. Sci. USA* **107,** 22399–22406 (2010).
42. Shen, J., Tareste, D. C., Paumet, F., Rothman, J. E. & Melia, T. J. Selective activation of cognate SNAREpins by Sec1/Munc18 proteins. *Cell* **128,** 183–195 (2007).
43. Xu, Y., Su, L. & Rizo, J. Binding of Munc18–1 to synaptobrevin and to the SNARE four-helix bundle. *Biochemistry* **49,** 1568–1576 (2010).
44. Zilly, F. E., Sørensen, J. B., Jahn, R. & Lang, T. Munc18-bound syntaxin readily forms SNARE complexes with synaptobrevin in native plasma membranes. *PLoS Biol.* **4,** e330 (2006).
45. Wojcik, S. M. & Brose, N. Regulation of membrane fusion in synaptic excitation-secretion coupling: speed and accuracy matter. *Neuron* **55,** 11–24 (2007).
46. Basu, J. *et al.* A minimal domain responsible for Munc13 activity. *Nature Struct. Mol. Biol.* **12,** 1017–1018 (2005).
47. Koch, H., Hofmann, K. & Brose, N. Definition of Munc13-homology-domains and characterization of a novel ubiquitously expressed Munc13 isoform. *Biochem. J.* **349,** 247–253 (2000).
48. Li, W. *et al.* The crystal structure of a Munc13 C-terminal module exhibits a remarkable similarity to vesicle tethering factors. *Structure* **19,** 1443–1455 (2011).
This crystal structure demonstrates that Munc13 is a member of the conserved CATCHR protein family involved in vesicle tethering.
49. Hammarlund, M., Palfreyman, M. T., Watanabe, S., Olsen, S. & Jorgensen, E. M. Open syntaxin docks synaptic vesicles. *PLoS Biol.* **5,** e198 (2007).
50. Richmond, J. E., Weimer, R. M. & Jorgensen, E. M. An open form of syntaxin bypasses the requirement for UNC-13 in vesicle priming. *Nature* **412,** 338–341 (2001).
51. Koushika, S. P. *et al.* A post-docking role for active zone protein Rim. *Nature Neurosci.* **4,** 997–1005 (2001).
52. Kochubey, O., Lou, X. & Schneggenburger, R. Regulation of transmitter release by Ca^{2+} and synaptotagmin: insights from a large CNS synapse. *Trends Neurosci.* **34,** 237–246 (2011).
53. Pang, Z. P. & Südhof, T. C. Cell biology of Ca^{2+}-triggered exocytosis. *Curr. Opin. Cell Biol.* **22,** 496–505 (2010).
54. Chapman, E. R. How does synaptotagmin trigger neurotransmitter release? *Annu. Rev. Biochem.* **77,** 615–641 (2008).
55. Groffen, A. J. *et al.* Doc2b is a high-affinity Ca^{2+} sensor for spontaneous neurotransmitter release. *Science* **327,** 1614–1618 (2010).
56. Yao, J., Gaffaney, J. D., Kwon, S. E. & Chapman, E. R. Doc2 is a Ca^{2+} sensor required for asynchronous neurotransmitter release. *Cell* **147,** 666–677 (2011).
57. Walter, A. M., Groffen, A. J., Sørensen, J. B. & Verhage, M. Multiple Ca^{2+} sensors in secretion: teammates, competitors or autocrats? *Trends Neurosci.* **34,** 487–497 (2011).

58. Xue, M. *et al.* Binding of the complexin N terminus to the SNARE complex potentiates synaptic-vesicle fusogenicity. *Nature Struct. Mol. Biol.* **17,** 568–575 (2010).
59. Yang, X., Kaeser-Woo, Y. J., Pang, Z. P., Xu, W. & Südhof, T. C. Complexin clamps asynchronous release by blocking a secondary Ca^{2+} sensor via its accessory alpha helix. *Neuron* **68,** 907–920 (2010).
60. Lai, A. L., Huang, H., Herrick, D. Z., Epp, N. & Cafiso, D. S. Synaptotagmin 1 and SNAREs form a complex that is structurally heterogeneous. *J. Mol. Biol.* **405,** 696–706 (2011).
61. Vrljic, M. *et al.* Molecular mechanism of the synaptotagmin-SNARE interaction in Ca^{2+}-triggered vesicle fusion. *Nature Struct. Mol. Biol.* **17,** 325–331 (2010).
62. Rizo, J., Chen, X. & Arac, D. Unraveling the mechanisms of synaptotagmin and SNARE function in neurotransmitter release. *Trends Cell Biol.* **16,** 339–350 (2006).
63. Chen, X. *et al.* Three-dimensional structure of the complexin/SNARE complex. *Neuron* **33,** 397–409 (2002).
64. Bracher, A., Kadlec, J., Betz, H. & Weissenhorn, W. X-ray structure of a neuronal complexin-SNARE complex from squid. *J. Biol. Chem.* **277,** 26517–26523 (2002).
65. Brose, N. For better or for worse: complexins regulate SNARE function and vesicle fusion. *Traffic* **9,** 1403–1413 (2008).
66. Stein, A. & Jahn, R. Complexins living up to their name–new light on their role in exocytosis. *Neuron* **64,** 295–297 (2009).
67. Neher, E. Complexin: does it deserve its name? *Neuron* **68,** 803–806 (2010).
68. Kummel, D. *et al.* Complexin cross-links prefusion SNAREs into a zigzag array. *Nature Struct. Mol. Biol.* **18,** 927–933 (2011).
69. Chicka, M. C., Hui, E., Liu, H. & Chapman, E. R. Synaptotagmin arrests the SNARE complex before triggering fast, efficient membrane fusion in response to Ca^{2+}. *Nature Struct. Mol. Biol.* **15,** 827–835 (2008).
70. Stein, A., Radhakrishnan, A., Riedel, D., Fasshauer, D. & Jahn, R. Synaptotagmin activates membrane fusion through a Ca^{2+}-dependent *trans* interaction with phospholipids. *Nature Struct. Mol. Biol.* **14,** 904–911 (2007).
71. Xue, M., Ma, C., Craig, T. K., Rosenmund, C. & Rizo, J. The Janus-faced nature of the C$_2$B domain is fundamental for synaptotagmin-1 function. *Nature Struct. Mol. Biol.* **15,** 1160–1168 (2008).
72. Lee, H. K. *et al.* Dynamic Ca^{2+}-dependent stimulation of vesicle fusion by membrane-anchored synaptotagmin 1. *Science* **328,** 760–763 (2010).
73. Martens, S., Kozlov, M. M. & McMahon, H. T. How synaptotagmin promotes membrane fusion. *Science* **316,** 1205–1208 (2007).
74. Hui, E., Johnson, C. P., Yao, J., Dunning, F. M. & Chapman, E. R. Synaptotagmin-mediated bending of the target membrane is a critical step in Ca^{2+}-regulated fusion. *Cell* **138,** 709–721 (2009).
75. Araç, D. *et al.* Close membrane-membrane proximity induced by Ca^{2+}-dependent multivalent binding of synaptotagmin-1 to phospholipids. *Nature Struct. Mol. Biol.* **13,** 209–217 (2006).
This study shows that synaptotagmin binds simultaneously to two membranes, bringing them into close proximity.
76. Sørensen, J. B. *et al.* Sequential N- to C-terminal SNARE complex assembly drives priming and fusion of secretory vesicles. *EMBO J.* **25,** 955–966 (2006).
77. Walter, A. M., Wiederhold, K., Bruns, D., Fasshauer, D. & Sørensen, J. B. Synaptobrevin N-terminally bound to syntaxin-SNAP-25 defines the primed vesicle state in regulated exocytosis. *J. Cell Biol.* **188,** 401–413 (2010).
78. Wiederhold, K. *et al.* A coiled coil trigger site is essential for rapid binding of synaptobrevin to the SNARE acceptor complex. *J. Biol. Chem.* **285,** 21549–21559 (2010).
79. Diao, J., Ishitsuka, Y. & Bae, W. R. Single-molecule FRET study of SNARE-mediated membrane fusion. *Biosci. Rep.* **31,** 457–463 (2011).
80. van den Bogaart, G. *et al.* One SNARE complex is sufficient for membrane fusion. *Nature Struct. Mol. Biol.* **17,** 358–364 (2010).
81. Chernomordik, L. V. & Zimmerberg, J. Bending membranes to the task: structural intermediates in bilayer fusion. *Curr. Opin. Struct. Biol.* **5,** 541–547 (1995).
82. McMahon, H. T., Kozlov, M. M. & Martens, S. Membrane curvature in synaptic vesicle fusion and beyond. *Cell* **140,** 601–605 (2010).
83. van den Bogaart, G. *et al.* Synaptotagmin-1 may be a distance regulator acting upstream of SNARE nucleation. *Nature Struct. Mol. Biol.* **18,** 805–812 (2011).
84. Kuo, W., Herrick, D. Z. & Cafiso, D. S. Phosphatidylinositol 4,5-bisphosphate alters synaptotagmin 1 membrane docking and drives opposing bilayers closer together. *Biochemistry* **50,** 2633–2641 (2011).
85. Schwartz, M. L. & Merz, A. J. Capture and release of partially zipped trans-SNARE complexes on intact organelles. *J. Cell Biol.* **185,** 535–549 (2009).
86. Rosenmund, C. & Stevens, C. F. Definition of the readily releasable pool of vesicles at hippocampal synapses. *Neuron* **16,** 1197–1207 (1996).
87. Pabst, S. *et al.* Rapid and selective binding to the synaptic SNARE complex suggests a modulatory role of complexins in neuroexocytosis. *J. Biol. Chem.* **277,** 7838–7848 (2002).
88. Mohrmann, R., de Wit, H., Verhage, M., Neher, E. & Sørensen, J. B. Fast vesicle fusion in living cells requires at least three SNARE complexes. *Science* **330,** 502–505 (2010).
Using a titration approach this study and ref. 89 reveal that neurotransmitter release requires only few SNARE complexes.
89. Sinha, R., Ahmed, S., Jahn, R. & Klingauf, J. Two synaptobrevin molecules are sufficient for vesicle fusion in central nervous system synapses. *Proc. Natl Acad. Sci. USA* **108,** 14318–14323 (2011).
90. Shi, L. *et al.* SNARE proteins: one to fuse and three to keep the nascent fusion pore open. *Science* **335,** 1355–1359 (2012).
91. Chan, Y. H., van Lengerich, B. & Boxer, S. G. Effects of linker sequences on vesicle fusion mediated by lipid-anchored DNA oligonucleotides. *Proc. Natl Acad. Sci. USA* **106,** 979–984 (2009).
92. Simonsson, L., Jonsson, P., Stengel, G. & Hook, F. Site-specific DNA-controlled fusion of single lipid vesicles to supported lipid bilayers. *ChemPhysChem* **11,** 1011–1017 (2010).
93. Lygina, A. S., Meyenberg, K., Jahn, R. & Diederichsen, U. Transmembrane domain peptide/peptide nucleic acid hybrid as a model of a SNARE protein in vesicle fusion. *Angew. Chem. Int. Edn Engl.* **50,** 8597–8601 (2011).
94. Robson Marsden, H., Elbers, N. A., Bomans, P. H., Sommerdijk, N. A. & Kros, A. A reduced SNARE model for membrane fusion. *Angew. Chem. Int. Edn Engl.* **48,** 2330–2333 (2009).
95. Avinoam, O. & Podbilewicz, B. Eukaryotic cell–cell fusion families. *Curr. Top. Membr.* **68,** 209–234 (2011).
96. Harrison, S. C. Viral membrane fusion. *Nature Struct. Mol. Biol.* **15,** 690–698 (2008).
97. Westermann, B. Mitochondrial fusion and fission in cell life and death. *Nature Rev. Mol. Cell Biol.* **11,** 872–884 (2010).
98. Moss, T. J., Daga, A. & McNew, J. A. Fusing a lasting relationship between ER tubules. *Trends Cell Biol.* **21,** 416–423 (2011).
99. Kozlov, M. M. & Markin, V. S. Possible mechanism of membrane fusion [in Russian]. *Biofizika* **28,** 242–247 (1983).
100. Yang, L. & Huang, H. W. Observation of a membrane fusion intermediate structure. *Science* **297,** 1877–1879 (2002).
101. Aeffner, S., Reusch, T., Weinhausen, B. & Salditt, T. Structure, hydration barrier and curvature of membrane hemifusion stalks with varying lipid composition obtained by x-ray diffraction. *Proc. Natl Acad. Sci. USA.* (in the press).
102. Qian, S. & Huang, H. W. A novel phase of compressed bilayers that models the prestalk transition state of membrane fusion. *Biophys. J.* **102,** 48–55 (2012).
103. Chernomordik, L. V. & Kozlov, M. M. Protein-lipid interplay in fusion and fission of biological membranes. *Annu. Rev. Biochem.* **72,** 175–207 (2003).
104. Kinnunen, P. K. Fusion of lipid bilayers: a model involving mechanistic connection to HII phase forming lipids. *Chem. Phys. Lipids* **63,** 251–258 (1992).
105. Risselada, H. J. & Grubmuller, H. How SNARE molecules mediate membrane fusion: recent insights from molecular simulations. *Curr. Opin. Struct. Biol.* **22,** 187–196 (2012).
106. Kozlov, M. M., McMahon, H. T. & Chernomordik, L. V. Protein-driven membrane stresses in fusion and fission. *Trends Biochem. Sci.* **35,** 699–706 (2010).
107. Kesavan, J., Borisovska, M. & Bruns, D. v-SNARE actions during Ca^{2+}-triggered exocytosis. *Cell* **131,** 351–363 (2007).
This study systematically measures the effect of extending the juxtamembrane region of synaptobrevin on neurotransmitter release.
108. Knecht, V. & Grubmuller, H. Mechanical coupling via the membrane fusion SNARE protein syntaxin 1A: a molecular dynamics study. *Biophys. J.* **84,** 1527–1547 (2003).
109. Risselada, H. J., Kutzner, C. & Grubmuller, H. Caught in the act: visualization of SNARE-mediated fusion events in molecular detail. *ChemBioChem* **12,** 1049–1055 (2011).
Using coarse-grain simulations, the transition states involved in SNARE-mediated membrane fusion are described on the basis of first principles.
110. Xu, Y., Zhang, F., Su, Z., McNew, J. A. & Shin, Y. K. Hemifusion in SNARE-mediated membrane fusion. *Nature Struct. Mol. Biol.* **12,** 417–422 (2005).
111. Wang, T., Smith, E. A., Chapman, E. R. & Weisshaar, J. C. Lipid mixing and content release in single-vesicle, SNARE-driven fusion assay with 1–5 ms resolution. *Biophys. J.* **96,** 4122–4131 (2009).
112. Giraudo, C. G. *et al.* SNAREs can promote complete fusion and hemifusion as alternative outcomes. *J. Cell Biol.* **170,** 249–260 (2005).
113. Reese, C., Heise, F. & Mayer, A. Trans-SNARE pairing can precede a hemifusion intermediate in intracellular membrane fusion. *Nature* **436,** 410–414 (2005).
114. Chernomordik, L. V. & Kozlov, M. M. Membrane hemifusion: crossing a chasm in two leaps. *Cell* **123,** 375–382 (2005).
115. Chernomordik, L. V. & Kozlov, M. M. Mechanics of membrane fusion. *Nature Struct. Mol. Biol.* **15,** 675–683 (2008).
116. Laage, R., Rohde, J., Brosig, B. & Langosch, D. A conserved membrane-spanning amino acid motif drives homomeric and supports heteromeric assembly of presynaptic SNARE proteins. *J. Biol. Chem.* **275,** 17481–17487 (2000).
117. Fuson, K. L., Montes, M., Robert, J. J. & Sutton, R. B. Structure of human synaptotagmin 1 C2AB in the absence of Ca^{2+} reveals a novel domain association. *Biochemistry* **46,** 13041–13048 (2007).

Acknowledgements Work in the authors' laboratories was supported by grants from the National Institutes of Health (3P01GM072694-05S1) to D.F. and R.J., of the Swiss National Fond to D.F. (31003A_133055) and of the Deutsche Forschungsgemeinschaft to D.F. (FA 297/3-1) and R.J. (SFB 803). The authors thank H. Grubmüller, E. Neher, J. Rissellada, G. van den Bogaart, M. Hernandez, J. Sørensen and J. Rizo for discussions and critical reading of the manuscript. We apologize to all colleagues whose work, although relevant, could not be mentioned and/or cited owing to space limitations.

Author Contributions Both authors wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence should be addressed to R.J. (rjahn@gwdg.de).

doi:10.1038/nature11518

The hexadehydro-Diels-Alder reaction

Thomas R. Hoye[1], Beeraiah Baire[1], Dawen Niu[1], Patrick H. Willoughby[1] & Brian P. Woods[1]

Arynes (aromatic systems containing, formally, a carbon–carbon triple bond) are among the most versatile of all reactive intermediates in organic chemistry. They can be 'trapped' to give products that are used as pharmaceuticals, agrochemicals, dyes, polymers and other fine chemicals. Here we explore a strategy that unites the *de novo* generation of benzynes—through a hexadehydro-Diels–Alder reaction—with their *in situ* elaboration into structurally complex benzenoid products. In the hexadehydro-Diels–Alder reaction, a 1,3-diyne is engaged in a [4+2] cycloisomerization with a 'diynophile' to produce the highly reactive benzyne intermediate. The reaction conditions for this simple, thermal transformation are notable for being free of metals and reagents. The subsequent and highly efficient trapping reactions increase the power of the overall process. Finally, we provide examples of how this *de novo* benzyne generation approach allows new modes of intrinsic reactivity to be revealed.

ortho-Benzyne (1,2-didehydrobenzene, **1**, Fig. 1a) is one of the oldest[1], most interesting, most useful and most well-studied of all reactive intermediates in chemistry. The multifaceted and often remarkably efficient reactions of benzynes with suitable trapping reagents (**1**→**3** in Fig. 1a) have long been employed in the service of synthetic chemistry. Even by 1967 many such reactions were known[2]. Nonetheless, newly discovered benzyne reaction motifs continue to emerge, especially so within the last decade[3–10]. This renaissance attests to the additional versatility and heralds even greater potential of this intermediate. Currently, all practical methods for producing benzynes involve the removal of two adjacent atoms or substituents ('G' and 'X'; Fig. 1a) from benzenoid precursors (**2**). These protocols typically require the use of a strong base and/or access to a 1,2-disubstituted arene substrate. A complementary general method for benzyne/aryne generation would considerably expand the preparative utility of these remarkable intermediates.

The venerable Diels–Alder cycloaddition reaction[11] is highly regarded in synthetic chemistry[12,13]. The prototypical event (Fig. 1b), found in every introductory organic chemistry textbook, is the combination of 1,3-butadiene (**5**) as the 4π-component with ethylene (**4**) as the dienophile to give cyclohexene (**6**)—a product in the tetrahydrobenzene oxidation state. If, instead, an alkyne like ethyne (**7**) is the dienophile, a 1,4-cyclohexadiene (here 1,4-dihydrobenzene (**8**)) results (Fig. 1c); we suggest this be viewed as a didehydro-Diels–Alder reaction. Another known variant (Fig. 1d) involves engagement of a (yet more highly oxidized) 1,3-enyne (for example, **9**) as the 4π-component with an alkyne (for example, **7**). It is interesting to note that the first example of this process (the thermal dimerization of phenylpropiolic acid)[14] was described 30 years before the initial report of Diels and Alder[11]. The intermediate cyclic allene **10** rapidly rearranges via a [1,5] hydrogen atom shift to benzene (**11**). This process has until now been called,

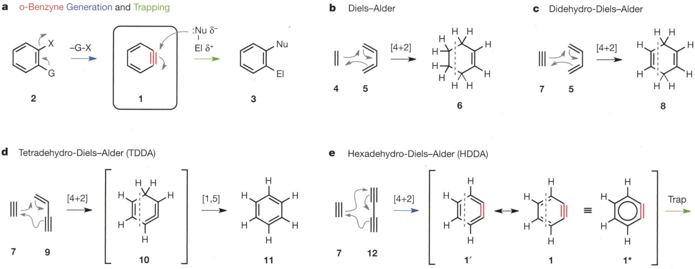


Figure 1 | Diels-Alder reactions of varying oxidation states. a, Generic benzyne generation (**2** to **1**) and trapping (**1** to **3**). Most commonly, G/X = H/halogen, CO_2^-/N_2^+, halogen/halogen, or TMS/OTf. **b–e**, Prototypes of Diels–Alder [4+2] (that is, the combination of a four-atom (or 4π electron) component with a two-atom (or 2π electron) component to produce a six-membered ring) reactions differing in the oxidation levels of the reactant pairs and products: classic Diels–Alder (**b**), didehydro-Diels–Alder (**c**), tetradehydro-Diels–Alder (TDDA) (**d**), and hexadehydro-Diels–Alder (HDDA, this work) (**e**) reactions. The aromatic character of benzyne is emphasized by the resonance contributors **1** and **1′** as well as by the resonance hybrid structure **1***. TMS, trimethylsilyl; OTf, trifluoromethanesulphonate; Nu, nucleophile; El, electrophile.

---

[1] Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA.

generically, the dehydro-Diels–Alder (DDA) reaction[15]. In the light of the generality of the results we present here, we suggest that the transformation in Fig. 1d would be better named the tetradehydro-Diels–Alder (TDDA) reaction.

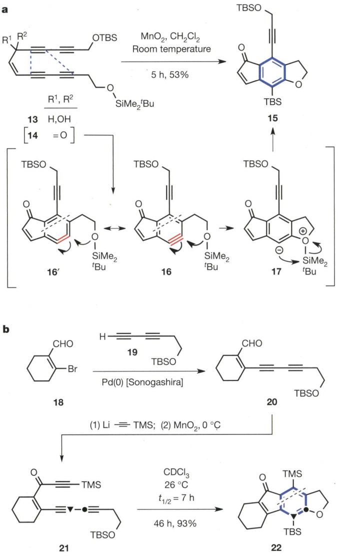


Figure 2 | Mechanistic rationale, substrate synthesis and mild conditions for our initial two HDDA reactions. a, Serendipitous observation of the HDDA reaction: cyclization of ketone **14** (see blue dashed lines) to putative benzyne (red) intermediate **16**/**16′** (via formation of the bonds indicated by the black dashed line in each) and subsequent trapping by the pendant silyl ether gave a hexasubstituted benzenoid (blue), the indenone **15**. **b**, Synthesis of **21** (via convergent coupling of **18** with **19**, addition of an ethynyl unit to **20**, and oxidation) and its facile, high-yielding conversion to the tetracyclic benzenoid (blue) **22**. The adjacent atoms denoted by the triangle and dot in substrate **21** map onto those of the product **22**. The half-life ($t_{1/2}$) for conversion of **21** to **22** was measured by *in situ* NMR analysis (Supplementary Fig. 1). TBS, *t*-butyldimethylsilyl.

The most highly oxidized Diels–Alder variant (Fig. 1e) is the cycloaddition between a 1,3-diyne like **12** and an alkyne diynophile like **7**, which generates *o*-benzyne (compare **1′**, **1** and **1***). This hexadehydro-Diels–Alder (HDDA) reaction is the subject of this Article. Given the efficiency, ease of precursor access, versatility and mild reaction conditions revealed by the examples we present here, it is remarkable that this reaction has remained essentially unexploited[16–19]. It is interesting to speculate that this may be due in part to the fact that the most common depiction of benzyne—the resonance contributor **1** (Fig. 1e)—obscures its potential construction via a [4+2] cycloaddition reaction. It is the alternative, but rarely encountered, Kekulé depiction (**1′**) that reveals the opportunity for assembly via an HDDA process.

We report below the broad scope of a strategy that combines the versatile and efficient generation of benzynes via the HDDA reaction with various trapping reactions to yield structurally complex benzenoid products. Each substrate is a readily accessible conjugated diyne containing a remote alkynyl diynophile. These cycloisomerize in (a highly exergonic) [4+2] fashion to produce the reactive aryne intermediate. The examples demonstrate that the tandem benzyne forming/trapping sequence can be designed to proceed with excellent efficiency.

The HDDA reaction revealed

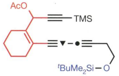
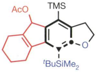
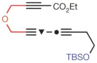
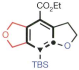
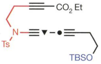
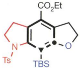
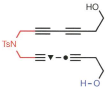
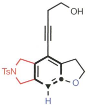
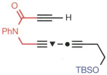
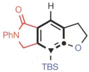
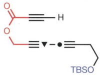
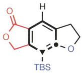
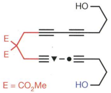
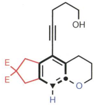
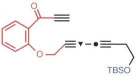
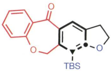
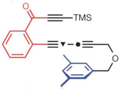
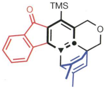
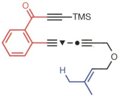
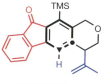
In the course of an otherwise unrelated study, we attempted to prepare the ketotetrayne **14** by oxidation of the precursor alcohol **13** with manganese dioxide (Fig. 2a). To our surprise, the major product from this experiment, formed in about 5 hours, was the (hexasubstituted) benzene derivative **15** (53% yield after purification). We quickly postulated that the benzyne intermediate **16′**/**16** was being both readily formed and efficiently trapped by the nucleophilic oxygen atom in the fortuitously poised β-siloxyethyl group. Migration of the silyl group from *O* to *C* within zwitterion **17**, a retro-Brook rearrangement, accounts for formation of **15**. This constitutes an unprecedented mode of benzyne trapping. Additionally, the process is attended by a substantial increase in structural complexity. The potential power of this transformation was immediately apparent.

We surmised that the modest yield observed in this reaction of tetrayne **14** reflected the fact that two competitive modes of [4+2] cyclization are possible. We were also keen to learn the feasibility of cyclization of analogous triynes. We therefore designed and synthesized a substrate—the ketotriyne **21** (Fig. 2b)—that could only undergo a HDDA reaction with a single regiochemical outcome. Our efforts were rewarded by its smooth transformation at room temperature to the hexasubstituted, tetracyclic indenone derivative **22** in 93% yield after chromatographic purification.

Intramolecular trapping

As the examples presented in Table 1 clearly show, the HDDA-initiated cascade has considerable scope with respect to both the cycloisomerization and the intramolecular trapping events. Each substrate is readily accessible by a convergent coupling strategy (compare Fig. 2b). All yields of purified products were ≥75%, and (with the exception of entry 8) all reactions occurred between temperatures of 20 and 120 °C. Highlights include: the presence of an electron withdrawing substituent on the diynophile enhances substrate reactivity (compare conditions for **21** to **22** (Fig. 2b) with entry 1); the activating carbonyl group can be a distal (carboalkoxy) rather than a tethering substituent (entries 2 and 3); many classical methods of aryne generation are not compatible with electron-withdrawing groups in the substrate[20]; carbonyl activation is not a necessity (entries 1, 4 and 7); products having nitrogen-containing heterocycles annulated to the new arene ring can be prepared (entries 3–5); an ester tether (entry 6) cyclizes more slowly than its *N*-phenyl amide analogue (entry 5), consistent with the lower concentration of the *s*-cis conformation required for ring closure; our observations are consistent with the absence of radical character in both the cycloaddition and trapping phases of the process (for example, reactions performed in chloroform solvent, an excellent hydrogen atom donor, have shown no evidence of hydrogen atom transfer (entries 4 and 9)); the new silyl ether trapping reaction has considerable generality (entries 1–3, 5, 6 and 8 and Fig. 2); other efficient internal benzyne traps include tethered alcohols (entries 4 and 7), aryl rings ([4+2] cycloaddition in entry 9), or alkenes (ene reaction in entry 10); seven-membered ring formation is feasible (entry 8), and the robust nature of the substrate and product at the high temperature required for this slower cyclization are noteworthy; and finally the silyl substituents in many of the products provide handles

Table 1 | Examples of intramolecular trapping of HDDA-generated benzynes

Entry	Substrate	Conditions	Product
1		110 °C 72 h PhMe 96%	
2		110 °C 20 h d_8-PhMe 86%	
3		120 °C 18 h PhMe 80%	
4		65 °C 20 h $CDCl_3$ 95%	
5		120 °C 15 h PhMe 92%	
6		120 °C 48 h PhMe 86%	
7		95 °C 48 h PhH 87%	
8		195 °C 32 h o-DCB 75%	
9		85 °C 18 h $CHCl_3$ 85%	
10		97 °C 22 h Heptane 83%	

Benzenoid (bold black bonds) synthesis via the HDDA cycloaddition has considerable substrate scope with respect to the nature of the poly-yne tether (red) and of the intramolecular trapping moiety (blue). Ac, acetyl; Ts, *p*-toluenesulphonyl; *o*-DCB, 1,2-dichlorobenzene.

for subsequent elaboration through protonative, oxidative or halogenative desilylation or cross-coupling reactions[21].

Intermolecular trapping

We wished to validate the feasibility of bimolecular trapping of these thermally generated benzynes. Clearly, this would add considerable versatility and power to a HDDA-initiated transformation. The triyne **23** (Fig. 3) represents a substrate that bears an innocent (non-participating) side chain. We have successfully captured the derived benzyne **24** (formed from **23** in a cyclization reaction having a half-life of ~4 h at about 85 °C) by a variety of external reagents to give adducts **25**. Highlights include: benzene as solvent forms the Diels–Alder adduct **25a** (compare entry 9 of Table 1)—although this process has been previously documented[22], because of the low reactivity of simple aromatics, rarely have they been trapped by benzynes in high yield; this result also indicates that many intramolecular trapping events are faster than capture by the aromatic solvents used in earlier examples (entries 1–3 and 5–8 of Table 1); the [2+2] cycloaddition of norbornene gives **25b** in higher yield than has been observed for trapping of arynes formed by conventional methods[2]; *N*-phenylacetamide gives **25c**, demonstrating that a nitrogenous substituent can be conveniently installed; acetic acid[22] or phenol traps **24** to cleanly provide adducts **25d** or **25e**, respectively, in processes that may share a mechanistic feature of transfer of a hydroxyl proton coincident with nucleophilic attack; this mode of reaction with acetic acid or phenol is unique and complementary to that seen with benzynes generated by non-reagent-free methods[23,24]; and finally net trapping by hydrogen bromide was achieved using $Br(CH_2)_2NH_2{\bullet}HBr$ in THF/H_2O (20:1) as an HBr source to give **25f** (6:1 mixture of isomers).

The sense of regioselectivity observed for formation of products **25c–f** is consistent with the analyses described in refs 24 and 25, in which the relative magnitude of the computed internal bond angles of an unsymmetrical benzyne is correlated with the site of nucleophilic attack. Namely, the more obtuse angle corresponds to the more electron deficient (δ^+) of the two benzyne carbon atoms. We computed the geometry for **24** ($R = CH_3$) using density functional theory (DFT) at the M06-2X/6-31+G(d,p)[26] level to have internal angles of 135° and 119° at atoms 'a' and 'b', respectively. We are currently investigating additional substrates that should allow us to distinguish the relative impact of electronic versus ring-distortion effects on the site of attack by external nucleophiles.

To gain understanding of some of the key thermodynamic features associated with the HDDA aryne-forming step, we turned to computational analysis of the reaction of ester **26** (Fig. 4). This simple triyne was cycloisomerized and the resulting benzyne trapped in *t*-butanol (120 °C, closed tube) to produce 5-*t*-butoxyphthalide (**28**) in 68% yield. Johnson and co-workers recently reported a DFT approach to compute the energetics of the hypothetical HDDA reaction of **7** with **12** to produce *o*-benzyne (**1**, Fig. 1e)[27]. They concluded that this benzyne-forming reaction was exothermic by $51.4\,\mathrm{kcal\,mol^{-1}}$. Using similar DFT methodology, we have computed the free energy of reaction for the conversion of triyne **26** to the aryne **27** and found it to be $-51\,\mathrm{kcal\,mol^{-1}}$. It is remarkable that highly reactive benzyne intermediates can be accessed by a thermal process that is exothermic by $\sim 50\,\mathrm{kcal\,mol^{-1}}$. These very favourable reaction energetics reflect the large amount of potential energy inherent in the (albeit kinetically protected) alkyne functional group. Finally, this point is further underscored by the computed free energy of reaction—namely, $-73\,\mathrm{kcal\,mol^{-1}}$—for the trapping by *t*-butanol of the strained alkyne[28] in **27**. Thus, the overall transformation of **26** to **28** is computed to be exothermic by $>120\,\mathrm{kcal\,mol^{-1}}$.

Discussion

Our results show that the HDDA reaction is a general method for generating benzynes under conditions amenable to a wide variety of intra- and intermolecular trapping events. This powerful and efficient

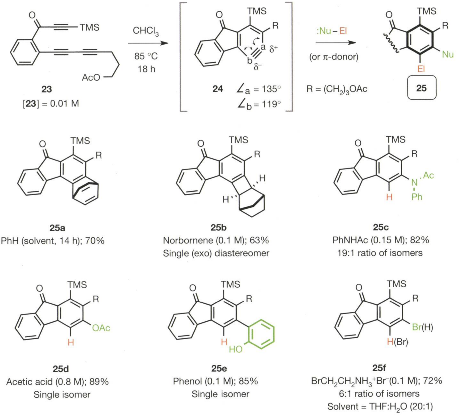

O
TMS
CHCl$_3$
85 °C
18 h
AcO
23
[23] = 0.01 M
24
∠a = 135°
∠b = 119°
δ+
δ−
R
:Nu–El
(or π-donor)
R = (CH$_2$)$_3$OAc
Nu
El
25
25a
PhH (solvent, 14 h); 70%
25b
Norbornene (0.1 M); 63%
Single (exo) diastereomer
H
25c
PhNHAc (0.15 M); 82%
19:1 ratio of isomers
N
Ac
Ph
25d
Acetic acid (0.8 M); 89%
Single isomer
OAc
25e
Phenol (0.1 M); 85%
Single isomer
HO
25f
BrCH$_2$CH$_2$NH$_3^+$Br$^-$ (0.1 M); 72%
6:1 ratio of isomers
Solvent = THF:H$_2$O (20:1)
Br(H)
H(Br)

Figure 3 | Examples of intermolecular trapping of HDDA-generated benzynes. Bimolecular trapping reactions of benzyne **24** to give adducts **25a–f.** Under the structure of each of these adducts is given information on the trapping agent (and amount) and yield following purification. Stereochemical information is given for adduct **25b**. Regiochemical information is given for adducts **25c–f**. THF, tetrahydrofuran.

domino sequence comprises a fundamentally new way to synthesize benzenoid compounds—especially those having high structural complexity. Because HDDA-derived benzynes are produced in the absence of by-products and external reagents, the intrinsic reactivity of a benzyne can be more fully explored and exploited. Even at this early stage of development, many additional lines of investigation have revealed themselves. These include studies of mechanistic issues (concerted versus stepwise cycloaddition), substituent effects (electronic versus steric), catalysis (Lewis acid or transition metal), new modes of benzyne trapping, the use of cleavable tethers, and the feasibility of bimolecular HDDA cycloadditions. This enabling technology is well-suited for preparing compound libraries, for accessing aromatics having substitution patterns that would be otherwise challenging to prepare, and for application to target molecule synthesis (for example, drugs, natural products[9,10], heteroaromatics and polyacenes). Lessons of how imagery (that is, **1** (versus **1′**) from **7** + **12**) can bias our perception are also embedded in the developments described here. Finally, our results serve as a reminder that, even in the twenty-first century, serendipitous discovery ("in the course of an otherwise unrelated study") in science can still play a pivotal role.

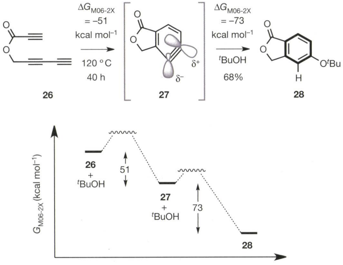

O
O
26
ΔG$_{\text{M06-2X}}$ = −51 kcal mol^{-1}
120 °C
40 h
27
δ+
δ−
ΔG$_{\text{M06-2X}}$ = −73 kcal mol^{-1}
^{t}BuOH
68%
28
O^{t}Bu
H
G$_{\text{M06-2X}}$ (kcal mol^{-1})
26 + ^{t}BuOH
51
27 + ^{t}BuOH
73
28

Figure 4 | Computed free energy changes for a representative HDDA-initiated cascade. Free energies of reaction ($\Delta G_{\text{M06-2X}}$ = $G_{\text{M06-2X}}$(product) − $G_{\text{M06-2X}}$(reactant)) were computed by DFT at the M06-2X/6-31+G(d,p) level for the aryne-generating (**26** to **27**) and *t*-butanol (^{t}BuOH) trapping (**27** to **28**) reactions. These serve to highlight the favourable thermodynamics associated with both the triyne cycloisomerization and aryne trapping events.

METHODS SUMMARY

General procedure. An oven-dried vial containing the triyne precursor in the indicated solvent (0.05 M) was closed with a Teflon-lined cap and heated at the indicated (external bath) temperature. After the indicated time the product was purified by chromatography on silica gel. Full procedural details and characterization data for all new compounds (substrates and products) and a detailed description of the computational protocols and results are provided in Supplementary Information.

Received 15 May; accepted 16 August 2012.

1. Wenk, H. H., Winkler, M. & Sander, W. One century of aryne chemistry. *Angew. Chem. Int. Edn* **42,** 502–528 (2003).
2. Hoffmann, R. W. *Dehydrobenzene and Cycloalkynes* (Organic Chemistry, A Series of Monographs, Vol. 11, Academic, 1967).
3. Pellissier, H. & Santelli, M. The use of arynes in organic synthesis. *Tetrahedron* **59,** 701–730 (2003).
4. Dyke, A. M., Hester, A. J. & Lloyd-Jones, G. C. Organometallic generation and capture of *ortho*-arynes. *Synthesis* 4093–4112 (2006).
5. Sanz, R. Recent applications of aryne chemistry to organic synthesis. A review. *Org. Prep. Proced. Int.* **40,** 215–291 (2008).

6. Gilchrist, T. L. in *Science of Synthesis* Vol. 43 (ed. Hopf, H.) 151–215 (Georg Thieme, 2008).
7. Chen, Y. & Larock, R. C. in *Modern Arylation Methods* (ed. Ackermann, L.) 401–473 (Wiley-VCH, 2009).
8. Kitamura, T. Synthetic methods for the generation and preparative application of benzyne. *Aust. J. Chem.* **63,** 987–1001 (2010).
9. Tadross, P. M. & Stoltz, B. M. A comprehensive history of arynes in natural product total synthesis. *Chem. Rev.* **112,** 3550–3577 (2012).
10. Gampe, C. M. & Carreira, E. M. Arynes and cyclohexyne in natural product synthesis. *Angew. Chem. Int. Edn* **51,** 3766–3778 (2012).
11. Diels, O. & Alder, K. Syntheses in the hydroaromatic series [in German]. *Justus Liebigs Ann. Chem.* **460,** 98–122 (1928).
12. Onishchenko, A. S. *Diene Synthesis* (Israel Program for Scientific Translations, 1964).
13. Nicolaou, K. C., Snyder, S. A., Montagnon, T. & Vassilikogiannakis, G. The Diels–Alder reaction in total synthesis. *Angew. Chem. Int. Edn* **41,** 1668–1698 (2002).
14. Michael, A. & Bucher, J. E. Über die Einwirkung von Eissigsäureanhydrid auf Phenylpropiolsäure. *Chem. Zentrblt.* 731–733 (1898).
15. Wessig, P. & Müller, G. The dehydro-Diels–Alder reaction. *Chem. Rev.* **108,** 2051–2063 (2008).
16. Bradley, A. Z. & Johnson, R. P. Thermolysis of 1,3,8-nonatriyne: evidence for intramolecular [2+4] cycloaromatization to a benzyne intermediate. *J. Am. Chem. Soc.* **119,** 9917–9918 (1997).
17. Miyawaki, K., Suzuki, R., Kawano, T. & Ueda, I. Cycloaromatization of a non-conjugated polyenyne system: synthesis of 5*H*-benzo[d]fluoreno[3,2-b]pyrans via diradicals generated from 1-[2-{4-(2-alkoxymethylphenyl)butan-1,3-diynyl}]phenylpentan-2,4-diyn-l-ols and trapping evidence for the 1,2-didehydrobenzene diradical. *Tetrahedr. Lett.* **38,** 3943–3946 (1997).
18. Kimura, H., Torikai, K., Miyawaki, K. & Ueda, I. Scope of the thermal cyclization of nonconjugated ene-yne-nitrile system: a facile synthesis of cyanofluorenol derivatives. *Chem. Lett.* **37,** 662–663 (2008).
19. Tsui, J. A. & Sterenberg, B. T. A metal-templated 4 + 2 cycloaddition reaction of an alkyne and a diyne to form a 1,2-aryne. *Organometallics* **28,** 4906–4908 (2009).
20. Uchiyama, M. *et al.* Generation of functionalized asymmetric benzynes with TMP-zincates. Effects of ligands on selectivity and reactivity of zincates. *J. Am. Chem. Soc.* **124,** 8514–8515 (2002).
21. Chang, W.-T. *et al.* Cross-coupling with organosilicon compounds. *Org. React.* **75,** 213–746 (2011).
22. Stiles, M., Miller, R. G. & Burckhardt, U. Reactions of benzyne intermediates in non-basic media. *J. Am. Chem. Soc.* **85,** 1792–1797 (1963).
23. Liu, Z. & Larock, R. C. Facile *O*-arylation of phenols and carboxylic acids. *Org. Lett.* **6,** 99–102 (2004).
24. Cheong, P. H.-Y. *et al.* Indolyne and aryne distortions and nucleophilic regioselectivities. *J. Am. Chem. Soc.* **132,** 1267–1269 (2010).
25. Garr, A. N. *et al.* Experimental and theoretical investigations into the unusual regioselectivity of 4,5-, 5,6-, and 6,7-indole aryne cycloadditions. *Org. Lett.* **12,** 96–99 (2010).
26. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120,** 215–241 (2008).
27. Ajaz, A. *et al.* Concerted vs. stepwise mechanisms in dehydro-Diels–Alder reactions. *J. Org. Chem.* **76,** 9320–9328 (2011).
28. Johnson, R. P. & Daoust, K. J. Interconversions of cyclobutyne, cyclopentyne, cyclohexyne, and their corresponding cycloalkylidene carbenes. *J. Am. Chem. Soc.* **117,** 362–367 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements P.H.W. thanks the National Science Foundation for a graduate research fellowship. Financial support from the National Institutes of Health (GM65597 and CA76497) is acknowledged. This work was carried out in part using hardware and software provided by the University of Minnesota Supercomputing Institute.

Author Contributions B.B. performed the initial experiments. All authors designed the experiments, analysed the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.R.H. (hoye@umn.edu).

doi:10.1038/nature11417

Small heat-shock proteins protect from heat-stroke-associated neurodegeneration

Nikos Kourtis[1], Vassiliki Nikoletopoulou[1] & Nektarios Tavernarakis[1]

Heat stroke is a life-threatening condition, characterized by catastrophic collapse of thermoregulation and extreme hyperthermia. In recent years, intensification of heat waves has caused a surge of heat-stroke fatalities. The mechanisms underlying heat-related pathology are poorly understood. Here we show that heat stroke triggers pervasive necrotic cell death and neurodegeneration in *Caenorhabditis elegans*. Preconditioning of animals at a mildly elevated temperature strongly protects from heat-induced necrosis. The heat-shock transcription factor HSF-1 and the small heat-shock protein HSP-16.1 mediate cytoprotection by preconditioning. HSP-16.1 localizes to the Golgi, where it functions with the Ca^{2+}- and Mn^{2+}-transporting ATPase PMR-1 to maintain Ca^{2+} homeostasis under heat stroke. Preconditioning also suppresses cell death inflicted by diverse insults, and protects mammalian neurons from heat cytotoxicity. These findings reveal an evolutionarily conserved mechanism that defends against diverse necrotic stimuli, and may be relevant to heat stroke and other pathological conditions involving necrosis in humans.

Heat-related pathologies such as heat stroke are estimated to soon become one of most serious causes of mortality[1]. Climatic shifts and other contemporary anthropogenic causes contribute to raise prevalence of hyperthermia incidents and associated deaths[2]. During heat stroke, core body temperature in excess of 40 °C elicits acute tissue injury and multi-organ failure that is often fatal[3]. The nervous system is particularly vulnerable. Heat-stroke survivors commonly suffer permanent neurological damage[4]. The devastating impact of hyperthermia on human health is a consequence of immediate effects of heat on cellular and organismal physiology, coupled with secondary inflammatory and coagulation responses[5]. Despite the severity and increasingly prevalent health risks associated with heat-inflicted damage, the cellular and molecular mechanisms responsible for the direct cytotoxicity of heat are not understood well.

Heat stroke kills by inducing necrosis

To gain insight into the molecular basis of heat cytotoxicity and to circumvent the confounding influence of secondary physiological and inflammatory responses, we developed and characterized a genetically tractable model of heat stroke in *C. elegans*, a poikilothermic, nematode worm that normally grows at ambient temperature. *C. elegans* is particularly suitable for investigating the direct cytotoxicity of heat because it lacks elaborate thermal regulation systems that would offset elevated external temperature by mounting homeothermic responses. Worms instantly equilibrate to ambient temperature. To simulate hyperthermia, we exposed synchronous nematode populations to 39 °C for 15 min. We then assessed animal survival 16 to 18 h after delivery of the noxious heat pulse (Supplementary Fig. 1a). Excessive heat caused immediate changes in animal behaviour and marked increase of mortality. Ultimately, less than 25% of animals survived heat stroke (Fig. 1a).

We observed widespread cell death in several tissues of afflicted individuals (Fig. 1b, c and Supplementary Fig. 2). Dying cells showed necrotic morphological features (Supplementary Fig. 1c and Supplementary Fig. 3a–e), expressed markers of necrotic death[6] (Supplementary Fig. 1d), and became permeable to propidium iodide (Supplementary Fig. 1e). Loss of key apoptosis or autophagy mediators did not suppress heat-stroke-induced mortality (Supplementary Fig. 4a, b). Moreover, autophagy was not upregulated in animals exposed to heat stroke (Supplementary Fig. 4c, d). Therefore, cell death consequent to heat stroke does not depend on the core apoptotic or autophagic machinery. In contrast, depletion of proteins required for necrosis[6,7] strongly promotes survival after heat stroke (Supplementary Fig. 4e). Thus, we conclude that heat stroke compromises viability by triggering extensive necrotic cell death in the nematode.

Preconditioning defends against hyperthermia

All organisms possess homeostatic mechanisms that promote survival by mounting adaptive responses to stressors[8]. Acute insults such as extreme heat may cause necrosis by overwhelming cellular homeostasis. Interestingly, pre-exposure to mild stress often elicits increased resistance to subsequent severe stress, a phenomenon termed hormesis[9]. We investigated whether experiencing intermediate, non-lethal temperature before heat stroke (preconditioning) evokes a similar hormetic effect in *C. elegans*. We found that preconditioning at 34 °C for 30 min (Supplementary Fig. 1a) markedly enhanced the capacity of animals to withstand heat stroke (Fig. 1a and Supplementary Figs 1c and 2). Survivors maintained viability over several days after heat stroke and showed no behavioural defects compared to untreated controls (Supplementary Fig. 5 and Supplementary Table 1).

Preconditioning may exert its protective effect by engaging the heat-shock-response pathway to fortify cells under extreme temperature. To test this idea, we examined mutants lacking HSF-1, the transcription factor that coordinates the heat-shock response[10]. Preconditioning does not increase survival of these animals, indicating that HSF-1 is required for protection against heat stroke. Moreover, overexpression of HSF-1 suppresses necrotic cell death and augments survival further, both after and without preconditioning (Fig. 1a and Supplementary Fig. 2). Therefore, HSF-1 is necessary and sufficient to render animals resistant to heat stroke.

Low insulin-like signalling (ILS) increases intrinsic thermotolerance of *C. elegans* during prolonged incubation at intermediate temperature (35 °C)[11]. We find that ILS deficiency enhances survival after heat stroke, both with and without preconditioning (Supplementary

---

[1] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology, Heraklion 71110, Crete, Greece.

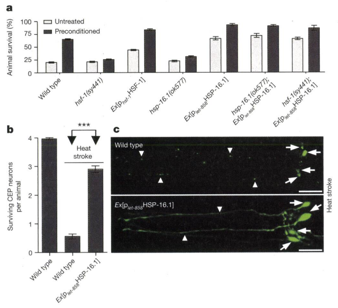


Figure 1 | Heat preconditioning protects against extreme thermal stress through HSF-1 and HSP-16.1. **a**, Survival of wild-type animals, *hsf-1(sy441)* or *hsp-16.1(ok577)* mutants, animals overexpressing *hsf-1* (*Ex*[$p_{hsf\text{-}1}$HSF-1]) or *hsp-16.1* (*Ex*[$p_{let\text{-}858}$HSP-16.1]), and *hsf-1* and *hsp-16.1* mutants overexpressing *hsp-16.1* after heat stroke, either without (untreated) or after preconditioning ($n = 350$ animals per assay; $P < 0.001$ for wild-type untreated versus preconditioned, $P > 0.05$ for *hsf-1(sy441)* or *hsp-16.1(ok577)* untreated versus preconditioned, $P < 0.001$ for wild-type untreated versus *Ex*[$p_{hsf\text{-}1}$HSF-1] or *Ex*[$p_{let\text{-}858}$HSP-16.1] untreated, $P < 0.001$ for *hsp-16.1(ok577)* untreated versus *hsp-16.1(ok577);Ex*[$p_{let\text{-}858}$HSP-16.1] untreated, $P < 0.001$ for *hsf-1(sy441)* untreated versus *hsf-1(sy441);Ex*[$p_{let\text{-}858}$HSP-16.1] untreated; two-way analysis of variance (ANOVA)). **b**, Survival of anterior CEP (cephalic) dopaminergic neurons of wild-type animals, and animals overexpressing *hsp-16.1* under heat stroke conditions ($n = 350$ animals per assay; $P < 0.001$ for wild-type heat-stroked versus *Ex*[$p_{let\text{-}858}$HSP-16.1] heat-stroked animals; unpaired *t*-test; $***P < 0.001$). Error bars, mean ± s.e.m. **c**, Images of the head region of wild-type animals (top panel) and animals overexpressing *hsp-16.1* (bottom panel) after heat stroke. Remnants of neuron cell bodies (arrows) and axonal beading (arrowheads) are seen in wild-type animals (top panel). Both soma (arrows) and axonal (arrowheads) architecture is preserved in animals overexpressing *hsp-16.1*. Scale bar, 20 μm.

Fig. 6a). HSF-1, which mediates intrinsic thermotolerance[12], is also required for the resistance of ILS mutants to heat stroke (Supplementary Fig. 6a). In addition to HSF-1, thermotolerance conferred by low ILS is mediated by the FOXO (forkhead box protein O) transcription factor DAF-16, which translocates to the nucleus under stress[13]. We found that DAF-16 accumulates in the nucleus after preconditioning (Supplementary Fig. 6b, c). Such relocalization of DAF-16 is consistent with a synergistic effect of HSF-1 and DAF-16 on heat-shock protein (HSP) gene promoters that facilitates their maximal expression[14]. Indeed, DAF-16 is partly required for protection against heat stroke by preconditioning or low ILS (Supplementary Fig. 6a).

We also tested the requirement for other key stress response regulators; SKN-1, which mediates oxidative stress resistance, and the hypoxia inducible factor HIF-1. After activation, SKN-1 translocates to the nucleus[15]. We did not detect nuclear accumulation of SKN-1 after preconditioning (Supplementary Fig. 6d). Moreover, SKN-1 is dispensable for cytoprotection by preconditioning or low ILS (Supplementary Fig. 6a, e). HIF-1 is also not needed for the protective effect of preconditioning (Supplementary Fig. 6f). Based on these findings, we conclude that HSF-1 specifically promotes survival under extreme thermal stress, following preconditioning at intermediate temperature.

After activation, HSF-1 induces expression of genes encoding HSPs, which protect cells from various cytotoxic conditions[16,17]. Based on their molecular mass, HSPs are classified into six main families[18]. A plausible mechanism pertinent to the strong shielding effect of preconditioning, involves the coordinated upregulation of batteries of heat-shock genes by HSF-1. To test this notion, we surveyed key representatives of all main *C. elegans* heat-shock-protein families for their role in heat-stress resistance acquired after preconditioning. Instead of a distributed contribution from several HSPs, we found that HSP-16.1, HSP-16.41 and DNJ-19 are primarily needed for acquired tolerance to heat stroke (Fig. 1a and Supplementary Fig. 7a). However, among the corresponding genes, only *hsp-16.1* and *hsp-16.41* show significant upregulation after preconditioning ($P < 0.001$ untreated versus preconditioned, unpaired *t*-test), whereas *dnj-19* is constitutively expressed, even under control conditions (Supplementary Fig. 7b). DNJ-19 is a co-chaperone that modulates the activity of HSP-70. We found that *hsp-70* expression is similarly elevated, and not subject to regulation by preconditioning (Supplementary Fig. 7b). Of the two *hsp* genes that are specifically induced by preconditioning, only overexpression of *hsp-16.1* conferred significant protection, bypassing the requirement for preconditioning to suppress necrosis and increase animal survival after heat stroke. Moreover, it further potentiated thermotolerance after preconditioning (Fig. 1 and Supplementary Figs 2 and 7c).

HSP-16.1 belongs to the family of α-crystallin domain-containing small heat shock proteins (sHSPs) that are strongly induced under heat stress or low insulin signalling, and contribute to stress resistance and longevity in *C. elegans*[14,19,20]. Transcriptional upregulation of *hsp-16.1* after preconditioning is fully dependent on HSF-1 (Supplementary Fig. 8a). In contrast, induction of *hsp-16.1* expression by low ILS is not abolished by loss of DAF-16 (Supplementary Fig. 8b).

Overexpression of *hsp-16.1* under the control of a heterologous, non-heat-shock-responsive promoter, bypasses the requirement for HSF-1 to defend against heat stroke (Fig. 1a). Our findings show that HSP-16.1 is specifically induced during preconditioning, and is both necessary and sufficient for resistance to hyperthermia. Other HSPs that seem also to be needed for extreme heat tolerance are either not specifically induced (for example, HSP-70 and DNJ-19) or are not sufficient for protection (HSP-16.41).

PMR-1 is required for hormetic protection

Typically, sHSPs are ubiquitously expressed and show widespread subcellular distribution[21]. Intriguingly, HSP-16.1 shows distinct localization to the medial Golgi of several cell types and is absent from other subcellular sites, including the closely related *trans*-Golgi and endosomes (Fig. 2a and Supplementary Figs 9a–d and 10a). HSP-16.41 displays a similar subcellular localization pattern (Supplementary Fig. 10b). This highly compartmentalized distribution is maintained under stress conditions in different genetic backgrounds (Supplementary Fig. 10c, d), suggesting that the two sHSPs function specifically in the medial Golgi to battle heat stress.

The Golgi apparatus is the main organelle for protein sorting and post-translational modification of proteins and lipids[22]. In addition, together with the endoplasmic reticulum and mitochondria, the Golgi is a major cellular Ca^{2+} reservoir, facilitating cellular ion homeostasis[23–25]. Perturbation of intracellular calcium concentration ($[Ca^{2+}]_i$) homeostasis has been implicated in necrotic cell death, both in mammals and *C. elegans*[26,27]. We considered whether heat stroke might inflict necrosis by diminishing the capacity of cells to maintain low $[Ca^{2+}]_i$ levels under noxious thermal stress. To test this hypothesis, we knocked down the *itr-1* gene encoding the inositol-1,4,5-trisphosphate receptor ($InsP_3R$) channel, which mediates Ca^{2+} trafficking out of both the Golgi and the endoplasmic reticulum[28]. $InsP_3R$ deficiency increased both intrinsic and acquired resistance to heat stroke (Fig. 2b). We found a similar effect after treatment of animals with dantrolene, a compound that inhibits Ca^{2+} release from

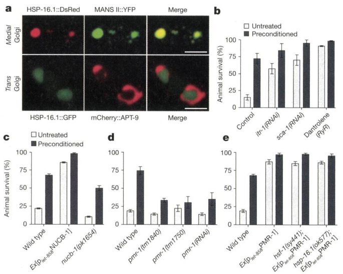


Figure 2 | HSP-16. 1 localizes to the medial Golgi and functions together with PMR-1 to mediate the protective effect of preconditioning against heat stroke. **a**, Images (cell bodies) of *C. elegans* neurons expressing fluorescently tagged HSP-16.1 along with medial (top row, middle panel) or *trans*-Golgi (bottom row, middle panel) markers. Scale bar, 6 μm. **b**, Survival of animals deficient for ITR-1 or SCA-1 (*itr-1(RNAi)*, *sca-1(RNAi)*), or animals exposed to dantrolene after heat stroke, either without (untreated) or after preconditioning ($n = 300$ animals per group; $P < 0.001$ for wild-type untreated (control) versus *itr-1(RNAi)*, *sca-1(RNAi)* or dantrolene-exposed; two-way ANOVA). **c**, Survival of wild-type animals, *nucb-1(pk1654)* mutants or animals overexpressing *nucb-1* (*Ex*[p$_{let\text{-}858}$NUCB-1]) after heat stroke, either without (untreated) or after preconditioning ($n = 300$ animals per assay; $P < 0.001$ for wild-type untreated versus *Ex*[p$_{let\text{-}858}$NUCB-1] untreated, $P < 0.01$ for wild-type untreated versus *nucb-1(pk1654)* untreated; two-way ANOVA). **d**, Survival of animals depleted for PMR-1 (*pmr-1(tm1840)* and *pmr-1(tm1750)* mutants, or *pmr-1(RNAi)* animals) under heat stroke, either without (untreated) or after preconditioning ($n = 800$ animals per assay; $P < 0.001$ for wild-type preconditioned versus *pmr-1(tm1840)*, *pmr-1(tm1750)* or *pmr-1(RNAi)* preconditioned; two-way ANOVA). **e**, Survival of wild-type animals and *hsf-1(sy441)* and *hsp-16.1(ok577)* mutants overexpressing *pmr-1* (*Ex*[p$_{let\text{-}858}$PMR-1]) under heat stroke, either without (untreated) or after heat preconditioning ($n = 300$ animals per assay; $P < 0.001$ for wild-type untreated versus *Ex*[p$_{let\text{-}858}$PMR-1], *hsf-1(sy441)*;*Ex*[p$_{let\text{-}858}$PMR-1] or *hsp-16.1(ok577)*;*Ex*[p$_{let\text{-}858}$PMR-1] untreated; two-way ANOVA). Error bars, mean ± s.e.m.

the Golgi and the endoplasmic reticulum, through the ryanodine receptor (RyR), or EGTA, a Ca^{2+} chelator (Fig. 2b and Supplementary Fig. 11). Hence, moderating $[Ca^{2+}]_i$ levels by blocking release of Golgi and endoplasmic reticulum Ca^{2+} stores, confers protection against thermal damage.

Ostensibly at odds with this idea, we observed a similar protective effect by knock down of *sca-1*, a gene encoding the sarco-endoplasmic reticulum Ca^{2+} ATPase (SERCA), which pumps Ca^{2+} from the cytoplasm into both the Golgi and the endoplasmic reticulum (Fig. 2b). Inhibition of SERCA induces endoplasmic reticulum stress and the unfolded protein response (UPR) pathway, leading to adaptation and survival[29,30]. We speculated that activation of the UPR^{ER} (UPR in the endoplasmic reticulum) may account for the paradoxical cytoprotection. To test this hypothesis, we monitored the UPR^{ER} after downregulation of SERCA, by means of an UPR^{ER} indicator[31]. SCA-1 depletion resulted in strong induction of UPR^{ER} (Supplementary Fig. 12a). We then tested whether elevated UPR^{ER} is sufficient to confer protection against heat stroke. Indeed, *xbp-1* mutants, which display marked constitutive increase of UPR^{ER} under physiological conditions[32], are resistant to heat stroke (Supplementary Fig. 12b). However, we found that preconditioning does not activate UPR^{ER} (Supplementary Fig. 13a, b), suggesting that induction of UPR^{ER} is not responsible for the protective effect of preconditioning. We also examined the involvement of mitochondrial UPR (UPR^{mt}), using a p$_{hsp\text{-}60}$GFP (green fluorescent protein) reporter[33]. Similar to the UPR^{ER}, the UPR^{mt} is not activated by preconditioning (Supplementary Fig. 13c, d). Our findings indicate that preconditioning, as implemented in our assay, is highly specific in activating solely the heat-shock-response pathway, without engaging other stress responses typically induced by prolonged incubation at elevated temperatures[34] (Supplementary Fig. 14a–d).

None of the above manipulations that render animals resistant to hyperthermia is specific to the Golgi, where HSP-16.1 localizes. To manipulate Ca^{2+} homeostasis specifically at the Golgi, we measured survival of transgenic animals overexpressing nucleobindin 1 (NUCB-1), a Golgi-resident Ca^{2+}-buffering protein[35], after heat stroke. We found that sequestration and retention of Ca^{2+} in the Golgi by nucleobindin is sufficient to protect from heat stroke, even without preconditioning. In contrast, animals deficient for NUCB-1 are sensitized to heat stroke-induced damage (Fig. 2c).

To further isolate and assess the contribution of the Golgi in Ca^{2+} homeostasis, we examined animals lacking PMR-1, a Golgi-specific Ca^{2+}- and Mn^{2+}-transporting P-type ATPase[36] that colocalizes with HSP-16.1 (Supplementary Fig. 15). PMR-1 deficiency suppresses preconditioning-acquired resistance to heat stroke (Fig. 2d). Importantly, overexpression of PMR-1 is sufficient to promote survival after heat stroke even without preconditioning, bypassing the requirement for both HSF-1 and HSP-16.1 (Fig. 2e). In addition, *pmr-1* expression is not altered by loss of HSF-1 or HSP-16.1 (Supplementary Fig. 16). Therefore, PMR-1 functions downstream of HSF-1 and HSP-16.1 to protect from hyperthermia. Together, our findings indicate that heat acclimatization through preconditioning may improve cytoprotection and animal survival by augmenting cellular Ca^{2+} homeostasis through the Golgi.

A universal cytoprotective mechanism

We considered the capacity of the mechanism involving specific heat-shock-response components and the Golgi PMR-1 ATPase to defend against additional necrotic insults unrelated to thermal stress. Dominant gain-of-function mutations in specific neuronal ion channels such as the degenerin MEC-4 and the acetylcholine receptor channel subunit DEG-3 evoke degenerative, necrotic cell death in *C. elegans* that is analogous to excitotoxicity in mammals[27,37]. Hypoxic conditions and overexpression of aggregation-prone proteins such as α-synuclein, also trigger necrosis in the nematode[38,39]. We exposed animals to brief heat preconditioning, before assaying cell death induced by toxic proteins or survival under hypoxia (Supplementary Fig. 1b). Remarkably, preconditioning suppressed neurodegeneration triggered by hyperactive ion channels or by overexpression of human α-synuclein (Fig. 3a, b and Supplementary Fig. 17a). Moreover, preconditioning increased animal resistance to hypoxic conditions (Fig. 3c). Preconditioning does not alter the expression, stability or function of cytotoxic proteins to ameliorate cell death (Supplementary Fig. 17b–d). In addition, suppression of neurodegeneration is not a mere consequence of a developmental delay in the onset of cell death (Supplementary Fig. 17e). Therefore, the heat-shock-response pathway intervenes downstream of necrotic insults to block necrosis. Cytoprotection is mediated by HSF-1 and HSP-16.1, which are both necessary and sufficient to ameliorate necrotic cell death (Fig. 3d, e). Thus, induction of the HSF-1–HSP-16 axis generally protects against necrosis inflicted by diverse stressors.

How might the heat-shock response neutralize necrotic stimuli? Excitotoxic insults initiate neuronal necrosis through elevation of $[Ca^{2+}]_i$ and subsequent aberrant activation of specific Ca^{2+}-dependent calpain proteases. In turn, calpains promote cell destruction by facilitating lysosome rupture[6,27,40]. We found that overexpression of NUCB-1 suppresses excitotoxic cell death, indicating that Ca^{2+} release from the Golgi contributes to neurodegeneration (Fig. 3f).

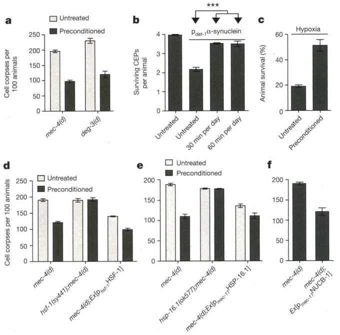


Figure 3 | HSF-1 and HSP-16.1 mediate protection against necrosis inflicted by diverse insults upon preconditioning. **a**, Number of neuron corpses at the L1 larval stage of development, per 100 animals carrying the neurotoxic dominant *(d)* *mec-4(d)* or *deg-3(d)* alleles, hatched from untreated or preconditioned eggs ($n = 350$ animals per assay; $P < 0.001$ for untreated versus preconditioned; two-way ANOVA). **b**, Dopaminergic neuron survival on the fifth day of adulthood in wild-type animals overexpressing α-synuclein, without or after receiving daily preconditioning (30 or 60 min per day; $n = 350$ animals per assay; $P < 0.001$ for untreated versus preconditioned; unpaired *t*-test; $***P < 0.001$). **c**, Survival of wild-type animals after near-lethal treatment with sodium azide ($\mathrm{NaN_3}$), either without or after preconditioning ($n = 350$ animals per assay; $P < 0.001$ for untreated versus preconditioned; unpaired *t*-test). **d**, Number of neuron corpses, at the L1 larval stage, per 100 *hsf-1(sy441);mec-4(d)* double mutants, and *mec-4(d)* animals overexpressing *hsf-1* hatched from untreated or preconditioned eggs ($n = 350$ animals per assay; $P > 0.05$ for *hsf-1(sy441);mec-4(d)* untreated versus preconditioned, $P < 0.001$ for *mec-4(d)* untreated versus *mec-4(d);Ex*[$\mathrm{p}_{hsf\text{-}1}$HSF-1]; two-way ANOVA). **e**, Number of neuron corpses at the L1 larval stage, per 100 *hsp-16.1(ok577);mec-4(d)* double mutants, and *mec-4(d)* animals overexpressing *hsp-16.1* hatched from untreated or preconditioned eggs ($n = 350$ animals per assay; $P > 0.05$ for *hsp-16.1(ok577);mec-4(d)* untreated versus preconditioned, $P < 0.001$ for *mec-4(d)* untreated versus *mec-4(d);Ex*[$\mathrm{p}_{mec\text{-}17}$HSP-16.1]; two-way ANOVA). **f**, Number of neuron corpses, at the L1 stage, per 100 *mec-4(d)* mutants overexpressing *nucb-1* ($n = 250$ animals; $P < 0.001$, unpaired *t*-test). Error bars, mean ± s.e.m.

To examine whether heat stroke also triggers neurodegeneration by increasing $[\mathrm{Ca}^{2+}]_i$, we monitored cytoplasmic Ca^{2+} levels in *C. elegans* neurons after heat stroke. Similar to the cytotoxic dominant MEC-4 ion channel, extreme thermal stress causes a sharp increase in $[\mathrm{Ca}^{2+}]_i$ levels (Fig. 4a, b). Heat preconditioning or inhibition of the $\mathrm{InsP_3R}$ channel, the SERCA pump and the RyR channel quenches $[\mathrm{Ca}^{2+}]_i$ elevation after subsequent heat stroke (Fig. 4a). Preconditioning by itself does not increase Ca^{2+} levels (Supplementary Fig. 18). These findings parallel the protective effect of the above interventions against heat stroke-induced damage and death (Fig. 2b). In contrast, loss of the Golgi-specific PMR-1 Ca^{2+} importer results in constitutive increase of $[\mathrm{Ca}^{2+}]_i$ levels (Fig. 4a), consistent with the requirement for PMR-1 in preconditioning-acquired thermotolerance (Fig. 2d). Moreover, loss of PMR-1 exacerbates degenerin-induced necrotic cell death (Supplementary Fig. 17f). Thus, Ca^{2+} import into the Golgi through PMR-1 is critical for survival after both heat stroke and excitotoxic insults. Combined, our data reveal a previously unsuspected and prominent role of the medial Golgi in maintaining Ca^{2+} homeostasis under cytotoxic stress.

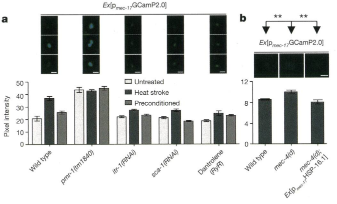


Figure 4 | Preconditioning requires PMR-1 to alleviate heat-stroke-induced cytoplasmic Ca^{2+} overload. **a**, Fluorescence intensity of neurons expressing the Ca^{2+} reporter GCaMP2.0, in wild-type animals, *pmr-1(tm1840)* mutants, and animals deficient for ITR-1, SCA-1 or treated with dantrolene, under normal conditions (untreated) and after heat stroke, either without or after preconditioning ($n = 60$ animals per assay; $P < 0.001$ for wild-type heat-stroked versus wild-type untreated or wild-type preconditioned; $P > 0.05$ for similar comparisons in mutant strains or after treatment with dantrolene; $P < 0.001$ for comparisons between *pmr-1(tm1840)* and other mutants or dantrolene treatment; two-way ANOVA). **b**, Fluorescence intensity of neurons, expressing GCaMP2.0, in wild-type animals, *mec-4(d)* mutants and *mec-4(d)* animals overexpressing *hsp-16.1* ($n = 150$ animals per assay; $P < 0.001$ for *mec-4(d)* versus wild-type or *mec-4(d);Ex*[$\mathrm{p}_{mec\text{-}17}$HSP-16.1]; unpaired *t*-test $**P < 0.01$). Representative images of neuron cell bodies are shown above each strain, for each condition. Scale bars, 10 μm. Error bars, mean ± s.e.m.

A conserved pathway against hyperthermia

Our findings using a *C. elegans* hyperthermia model implicate specific heat-shock-response components and Golgi-mediated Ca^{2+} homeostasis in resistance to a broad spectrum of necrotic insults, acquired after heat preconditioning. Has this protective mechanism been maintained during metazoan evolution or is it peculiar to the nematode? To address this question, we investigated whether key aspects of heat stroke neuropathology and heat preconditioning are conserved in mammalian neurons. To this end, we differentiated mouse embryonic stem cells into a homogeneous population of glutamatergic excitatory neurons[41]. Heat stroke induced pervasive death of embryonic-stem-cell-derived neurons (Fig. 5a, b and Supplementary Fig. 19a), and this effect could be largely prevented by heat preconditioning (Fig. 5a, b). We obtained similar results using primary cultures of mouse cortical (Fig. 5a, b) and striatal neurons (data not shown). In all cases, heat-stroked neurons showed axonal degeneration and other characteristics typical of necrotic death (Supplementary Fig. 19a, b).

Notably similar to the response in *C. elegans*, we found that overexpression of crystallin αA (CRYAA)[42], which co-localizes with the Golgi marker α-mannosidase II and the PMR1 ATPase (Supplementary Fig. 20a, b), was sufficient to protect mammalian neurons from heat-stroke-induced death, even in the absence of preconditioning (Fig. 5c). Finally, we tested whether the protective effect of preconditioning in mammalian neurons is also mediated by the PMR1 ATPase[43]. Heat stroke caused massive necrotic death and axonal degeneration in neurons expressing short hairpin RNAs (shRNAs) against *Pmr1*, even after preconditioning (Fig. 5d; Supplementary Fig. 21). Thus, knockdown of *Pmr1* abolishes the protective effect of preconditioning in mammalian neurons. These findings suggest that hormetic stress exposure engages a potent, evolutionarily conserved mechanism to defend against necrotic neurodegeneration, in metazoans as diverse as nematodes and mammals.

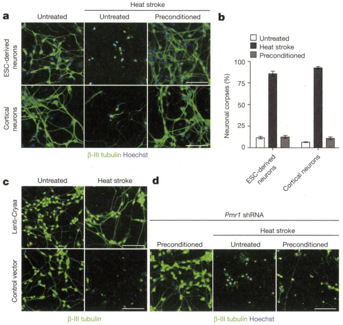


Figure 5 | Heat preconditioning protects mammalian neurons against extreme thermal stress through crystallin αA and PMR1. **a**, Images of embryonic stem cell (ESC)-derived neurons (top panels) and primary cortical neurons (bottom panels) immunostained for β-III tubulin, under normal conditions (untreated) and after heat stroke, either without or after preconditioning. Neuronal nuclei are stained with Hoechst. Scale bar, 100 μm. **b**, Corpses of ESC-derived and primary cortical neurons, under normal conditions (untreated) and after heat stroke, either without or after preconditioning ($n = 900$ neurons per assay, $P < 0.001$ for ESC-derived or cortical heat-stroked neurons versus corresponding preconditioned neurons; two-way ANOVA; error bars, mean ± s.e.m.). **c**, Images of ESC-derived neurons overexpressing *Cryaa*, or control vector after lentiviral infection, under normal conditions (untreated) and after heat stroke. Scale bar, 100 μm. **d**, Images of ESC-derived neurons after *Pmr1* downregulation, under normal conditions (untreated) and after heat stroke, either without or after preconditioning. Scale bar, 100 μm.

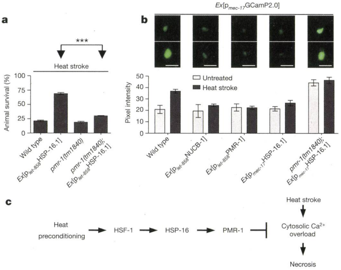


Figure 6 | HSP-16.1 requires PMR-1 to suppress heat-stroke-induced necrosis and cytoplasmic Ca^{2+} overload. **a**, Survival of wild-type animals or *pmr-1(tm1840)* mutants overexpressing *hsp-16.1*, under heat stroke ($n = 400$ animals per assay; $P < 0.001$ for $Ex[p_{let-858}$HSP-16.1] versus *pmr-1(tm1840)*;$Ex[p_{let-858}$HSP-16.1]; unpaired *t*-test; $^{***}P < 0.001$). **b**, Fluorescence intensity of touch receptor neurons expressing GCaMP2.0, in wild-type animals overexpressing *nucb-1*, *pmr-1* and *hsp-16.1*, or in *pmr-1(tm1840)* mutants overexpressing *hsp-16.1* under normal conditions (untreated), and after heat stroke ($n = 250$ animals per assay; $P < 0.001$ for wild-type untreated versus wild-type heat-stroked; $P > 0.05$ for $Ex[p_{let-858}$NUCB-1], $Ex[p_{let-858}$PMR-1], $Ex[p_{mec-17}$HSP-16.1] and *pmr-1(tm1840)*;$Ex[p_{mec-17}$HSP-16.1] untreated versus heat-stroked; two-way ANOVA). Error bars, mean ± s.e.m. Representative images of neuron cell bodies are depicted above each strain, for each condition. Scale bar, 10 μm. **c**, Schematic representation of the protective pathway by which preconditioning augments resistance to heat-stroke-induced necrotic cell death. Arrows denote activation or upregulation. Blunt-ended arrow denotes inhibition.

Discussion

In this report, we describe a versatile, genetically tractable *C. elegans* model of heat stroke and show that a single sHSP, HSP-16.1, is an effective general protector against multiple necrotic insults. HSP-16.1 localizes in the Golgi, where it functions together with the PMR-1 pump to prevent cytoplasmic Ca^{2+} overload under extreme stress. Hormetic induction of HSP-16.1 expression by brief heat preconditioning is sufficient to fortify cells against diverse insults. Our findings, in both *C. elegans* and mammalian neurons indicate that an evolutionarily conserved pathway underlies the protective response (Fig. 6c). Three specific predictions derive from this model. First, that PMR-1 is required for resistance to extreme temperature, conferred by HSP-16.1 overexpression. Second, that HSP-16.1, NUCB-1 and PMR-1, which are exclusively localized in the Golgi, may promote survival by augmenting Ca^{2+} sequestration into the Golgi and consequently maintaining subthreshold $[Ca^{2+}]_i$ levels. Third, importantly the model also predicts that moderation of $[Ca^{2+}]_i$ levels by HSP-16.1 requires PMR-1. We tested each of these predictions and found that, in complete agreement with the model, PMR-1 depletion diminishes the protective effect of HSP-16.1 overexpression (Fig. 6a). In addition, HSP-16.1, NUCB-1 or PMR-1 overexpression moderates $[Ca^{2+}]_i$ elevation induced by heat stroke (Fig. 6b). Moreover, the effect of HSP-16.1 is fully dependent on PMR-1 (Fig. 6b).

What is the molecular basis of HSP-16.1 protective function in the Golgi? sHSPs assemble into oligomeric complexes of variable stoichiometry and serve as molecular chaperones, efficiently binding denatured proteins and/or preventing irreversible protein aggregation and insolubilization[44]. Importantly, mammalian PMR1 is selectively impaired during ischaemic or reperfusion brain injury[45–47]. We propose that HSP-16.1 contributes to stabilize and protect the stress-labile PMR-1 pump, allowing for efficient clearance of Ca^{2+} from the cytoplasm, after necrotic insults (Supplementary Fig. 22a–d).

Our study reveals a novel protective mechanism operating specifically in the Golgi to defend against a wide range of necrosis initiators. Given the strong evolutionary conservation of the proteins involved, this mechanism is probably relevant to related human pathologies. Indeed, noxious heat disrupts cytoplasmic calcium Ca^{2+} absorption by subcellular organelles[48]. In addition, malignant hyperthermia susceptibility has been linked to mutations that cause resting Ca^{2+} leakage to the cytoplasm through RyR[49,50]. Therefore, our findings could facilitate identification of candidate common intervention targets, in the effort to battle necrosis-related disorders in humans.

METHODS SUMMARY

Nematode heat stroke and preconditioning assays. For heat stroke assays, synchronized populations of animals were grown under optimal growth conditions at 20 °C to young adult stage (early egg-laying). Synchronized populations were generated by transferring 20 to 30 gravid adult animals on fresh plates, letting them lay eggs for 1.5 h and then removing the adults. Animals were collected with M9 in 1.5-ml eppendorf tubes and incubated in 200 μl M9 buffer, at 39 °C for 15 min, in a water bath. Untreated worms were kept at 20 °C for the same time period. Survival rates were examined approximately 16 to 18 h post heat treatment. Animals that either moved freely or responded to tapping of the plate by moving the head or tail region of the body were considered survivors. To test whether prior activation of the heat-shock response (preconditioning) leads to increased survival against heat stroke, synchronized animals (prepared as described above) were incubated at 34 °C for 30 min, after a 20-min recovery incubation at 20 °C and then incubation at 39 °C for 15 min. Control animals were kept at 20 °C after the recovery incubation. Survival rates were examined approximately 16-18 h post heat treatment at 39 °C. Survivors were scored as above.

Full Methods and any associated references are available in the online version of the paper.

Received 30 March; accepted 16 July 2012.

Published online 12 September 2012.

1. Patz, J. A., Campbell-Lendrum, D., Holloway, T. & Foley, J. A. Impact of regional climate change on human health. *Nature* **438,** 310–317 (2005).
2. Rowlands, D. J. *et al.* Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nature Geosci.* **5,** 256–260 (2012).
3. Bouchama, A. & Knochel, J. P. Heat stroke. *N. Engl. J. Med.* **346,** 1978–1988 (2002).
4. Dhopesh, V. P. & Burns, R. A. Loss of nerve conduction in heat stroke. *N. Engl. J. Med.* **294,** 557–558 (1976).
5. Hall, D. M. *et al.* Mechanisms of circulatory and intestinal barrier dysfunction during whole body hyperthermia. *Am. J. Physiol. Heart Circ. Physiol.* **280,** H509–H521 (2001).
6. Syntichaki, P., Xu, K., Driscoll, M. & Tavernarakis, N. Specific aspartyl and calpain proteases are required for neurodegeneration in *C. elegans*. *Nature* **419,** 939–944 (2002).
7. Xu, K., Tavernarakis, N. & Driscoll, M. Necrotic cell death in *C. elegans* requires the function of calreticulin and regulators of Ca^{2+} release from the endoplasmic reticulum. *Neuron* **31,** 957–971 (2001).
8. Kourtis, N. & Tavernarakis, N. Cellular stress response pathways and ageing: intricate molecular relationships. *EMBO J.* **30,** 2520–2531 (2011).
9. Calabrese, E. J. Hormesis: a revolution in toxicology, risk assessment and medicine. *EMBO Rep.* **5,** S37–S40 (2004).
10. Åkerfelt, M., Morimoto, R. I. & Sistonen, L. Heat shock factors: integrators of cell stress, development and lifespan. *Nature Rev. Mol. Cell Biol.* **11,** 545–555 (2010).
11. McColl, G. *et al.* Insulin-like signaling determines survival during stress via posttranscriptional mechanisms in *C. elegans*. *Cell Metab.* **12,** 260–272 (2010).
12. Chiang, W. C., Ching, T. T., Lee, H. C., Mousigian, C. & Hsu, A. L. HSF-1 regulators DDL-1/2 link insulin-like signaling to heat-shock responses and modulation of longevity. *Cell* **148,** 322–334 (2012).
13. Lin, K., Hsin, H., Libina, N. & Kenyon, C. Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nature Genet.* **28,** 139–145 (2001).
14. Hsu, A. L., Murphy, C. T. & Kenyon, C. Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* **300,** 1142–1145 (2003).
15. An, J. H. & Blackwell, T. K. SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. *Genes Dev.* **17,** 1882–1893 (2003).
16. Lindquist, S. & Craig, E. A. The heat-shock proteins. *Annu. Rev. Genet.* **22,** 631–677 (1988).
17. Morimoto, R. I. Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes Dev.* **22,** 1427–1438 (2008).
18. Morimoto, R. I. Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes Dev.* **12,** 3788–3796 (1998).
19. Walker, G. A. & Lithgow, G. J. Lifespan extension in *C. elegans* by a molecular chaperone dependent upon insulin-like signals. *Aging Cell* **2,** 131–139 (2003).
20. Morley, J. F. & Morimoto, R. I. Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Mol. Biol. Cell* **15,** 657–664 (2004).
21. Haslbeck, M., Franzmann, T., Weinfurtner, D. & Buchner, J. Some like it hot: the structure and function of small heat-shock proteins. *Nature Struct. Mol. Biol.* **12,** 842–846 (2005).
22. Lowe, M. Structural organization of the Golgi apparatus. *Curr. Opin. Cell Biol.* **23,** 85–93 (2011).
23. Missiaen, L., Dode, L., Vanoevelen, J., Raeymaekers, L. & Wuytack, F. Calcium in the Golgi apparatus. *Cell Calcium* **41,** 405–416 (2007).
24. Van Baelen, K. *et al.* The Ca^{2+}/Mn^{2+} pumps in the Golgi apparatus. *Biochim. Biophys. Acta* **1742,** 103–112 (2004).
25. Pinton, P., Pozzan, T. & Rizzuto, R. The Golgi apparatus is an inositol 1,4,5-trisphosphate-sensitive Ca^{2+} store, with functional properties distinct from those of the endoplasmic reticulum. *EMBO J.* **17,** 5298–5308 (1998).
26. McCall, K. Genetic control of necrosis — another type of programmed cell death. *Curr. Opin. Cell Biol.* **22,** 882–888 (2010).
27. Syntichaki, P. & Tavernarakis, N. The biochemistry of neuronal necrosis: rogue biology? *Nature Rev. Neurosci.* **4,** 672–684 (2003).
28. Michelangeli, F., Ogunbayo, O. A. & Wootton, L. L. A plethora of interacting organellar Ca^{2+} stores. *Curr. Opin. Cell Biol.* **17,** 135–140 (2005).
29. Ferri, K. F. & Kroemer, G. Organelle-specific initiation of cell death pathways. *Nature Cell Biol.* **3,** E255–E263 (2001).
30. Rutkowski, D. T. *et al.* Adaptation to ER stress is mediated by differential stabilities of pro-survival and pro-apoptotic mRNAs and proteins. *PLoS Biol.* **4,** e374 (2006).
31. Calfon, M. *et al.* IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* **415,** 92–96 (2002).
32. Richardson, C. E., Kinkel, S. & Kim, D. H. Physiological IRE-1-XBP-1 and PEK-1 signaling in *Caenorhabditis elegans* larval development and immunity. *PLoS Genet.* **7,** e1002391 (2011).
33. Yoneda, T. *et al.* Compartment-specific perturbation of protein handling activates genes encoding mitochondrial chaperones. *J. Cell Sci.* **117,** 4055–4066 (2004).
34. Lithgow, G. J., White, T. M., Melov, S. & Johnson, T. E. Thermotolerance and extended life-span conferred by single-gene mutations and induced by thermal stress. *Proc. Natl Acad. Sci. USA* **92,** 7540–7544 (1995).
35. Lin, P., Yao, Y., Hofmeister, R., Tsien, R. Y. & Farquhar, M. G. Overexpression of CALNUC (nucleobindin) increases agonist and thapsigargin releasable Ca^{2+} storage in the Golgi. *J. Cell Biol.* **145,** 279–289 (1999).
36. Van Baelen, K., Vanoevelen, J., Missiaen, L., Raeymaekers, L. & Wuytack, F. The Golgi PMR1 P-type ATPase of *Caenorhabditis elegans*. Identification of the gene and demonstration of calcium and manganese transport. *J. Biol. Chem.* **276,** 10683–10691 (2001).
37. Yamashima, T. *et al.* Sustained calpain activation associated with lysosomal rupture executes necrosis of the postischemic CA1 neurons in primates. *Hippocampus* **13,** 791–800 (2003).
38. Cao, S., Gelwix, C. C., Caldwell, K. A. & Caldwell, G. A. Torsin-mediated protection from cellular stress in the dopaminergic neurons of *Caenorhabditis elegans*. *J. Neurosci.* **25,** 3801–3812 (2005).
39. Scott, B. A., Avidan, M. S. & Crowder, C. M. Regulation of hypoxic death in *C. elegans* by the insulin/IGF receptor homolog DAF-2. *Science* **296,** 2388–2391 (2002).
40. Yamashima, T. Implication of cysteine proteases calpain, cathepsin and caspase in ischemic neuronal death of primates. *Prog. Neurobiol.* **62,** 273–295 (2000).
41. Bibel, M., Richter, J., Lacroix, E. & Barde, Y. A. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nature Protocols* **2,** 1034–1043 (2007).
42. Brady, J. P. *et al.* Targeted disruption of the mouse αA-crystallin gene induces cataract and cytoplasmic inclusion bodies containing the small heat shock protein αB-crystallin. *Proc. Natl Acad. Sci. USA* **94,** 884–889 (1997).
43. Shull, G. E. *et al.* Physiological functions of plasma membrane and intracellular Ca^{2+} pumps revealed by analysis of null mutants. *Ann. NY Acad. Sci.* **986,** 453–460 (2003).
44. van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C. & Vierling, E. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Struct. Biol.* **8,** 1025–1030 (2001).
45. Gidday, J. M. Cerebral preconditioning and ischaemic tolerance. *Nature Rev. Neurosci.* **7,** 437–448 (2006).
46. Lehotský, J. *et al.* Ion transport systems as targets of free radicals during ischemia reperfusion injury. *Gen. Physiol. Biophys.* **21,** 31–37 (2002).
47. Pavlíková, M. *et al.* Alterations induced by ischemic preconditioning on secretory pathways Ca^{2+}-ATPase (SPCA) gene expression and oxidative damage after global cerebral ischemia/reperfusion in rats. *Cell. Mol. Neurobiol.* **29,** 909–916 (2009).
48. Greffrath, W., Kirschstein, T., Nawrath, H. & Treede, R. Changes in cytosolic calcium in response to noxious heat and their relationship to vanilloid receptors in rat dorsal root ganglion neurons. *Neuroscience* **104,** 539–550 (2001).
49. Lanner, J. T. e. t. a. l. AICAR prevents heat-induced sudden death in RyR1 mutant mice independent of AMPK activation. *Nature Med.* **18,** 244–251 (2012).
50. Protasi, F., Paolini, C. & Dainese, M. Calsequestrin-1: a new candidate gene for malignant hyperthermia and exertional/environmental heat stroke. *J. Physiol.* **587,** 3095–3100 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank K. Georgila, K. Palikaras and E. Bessa for help with cell-death assays, A. Pasparaki for technical support with experiments and C. Olendrowitz for help with electron microscopy. We thank S. Eimer for providing the Golgi and endosomal reporter constructs and for support with the electron microscopy, N. Chronis for providing the pN1-GCaMP2.0 plasmid, G. Caldwell for the α-synuclein-expressing *C. elegans* strain and K. Palikaras for the *skn-1* RNAi plasmid. We thank J. Vanoevelen and F. Wuytack for the antibody against PMR-1, S. Mitrovic for the antibody against mannosidase II, G. Sourvinos for lentiviral plasmids, and S. Gascon for the pLV plasmid. Some nematode strains used in this work were provided by the *Caenorhabditis* Genetics Center, which is funded by the National Center for Research Resources (NCRR) of the National Institutes of Health (NIH), and S. Mitani (National Bioresource Project) in Japan. We thank A. Fire for plasmid vectors. V.N. is supported by an European Molecular Biology Organization (EMBO) Long Term Fellowship. This work was funded by grants from the European Research Council (ERC) and the European Commission 7th Framework Programme.

Author Contributions N.K., V.N. and N.T. designed and carried out experiments. N.K. and N.T. analysed data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.T. (tavernarakis@imbb.forth.gr).

doi:10.1038/nature11529

Retinal waves coordinate patterned activity throughout the developing visual system

James B. Ackman[1], Timothy J. Burbridge[1] & Michael C. Crair[1]

The morphological and functional development of the vertebrate nervous system is initially governed by genetic factors and subsequently refined by neuronal activity. However, fundamental features of the nervous system emerge before sensory experience is possible. Thus, activity-dependent development occurring before the onset of experience must be driven by spontaneous activity, but the origin and nature of activity *in vivo* remains largely untested. Here we use optical methods to show in live neonatal mice that waves of spontaneous retinal activity are present and propagate throughout the entire visual system before eye opening. This patterned activity encompassed the visual field, relied on cholinergic neurotransmission, preferentially initiated in the binocular retina and exhibited spatiotemporal correlations between the two hemispheres. Retinal waves were the primary source of activity in the midbrain and primary visual cortex, but only modulated ongoing activity in secondary visual areas. Thus, spontaneous retinal activity is transmitted through the entire visual system and carries patterned information capable of guiding the activity-dependent development of complex intra- and inter-hemispheric circuits before the onset of vision.

The display of spontaneous activity is an emergent property of the immature nervous system that is thought to mediate synaptic competition[1] and instruct self-organization in many developing neural circuits[2–4]. If spontaneous activity is structured with respect to the topographic organization of first-order sensory and motor circuits, then these activity patterns could provide a template for activity-dependent development of downstream synaptic connectivity throughout the nervous system[5]. Whether spontaneous patterned activity exists and is communicated through all levels of organization for any sensory system during development is currently unknown.

In the developing visual system, isolated preparations of retina exhibit propagating bursts of action potentials among neighbouring retinal ganglion cells (RGCs), termed 'retinal waves'[6–8]. As RGCs relay visual information to higher-order structures in the central nervous system, retinal waves are thought to have a key role in the activity-dependent refinement of topographic neural maps in the superior colliculus, dorsal lateral geniculate nucleus (dLGN) and visual cortex[9–13], which exhibit functional connectivity before the onset of visual experience. However, the role of retinal waves in neural-circuit development remains controversial[14,15], partly because their existence has never been demonstrated *in vivo*.

Previous work using extracellular microelectrode recording techniques *in vivo* showed limited and local correlated spiking between pairs of embryonic rat RGCs[16], but no direct assessment of wave activity has been conducted *in vivo*, probably because of the methodological challenges associated with recording from a large cohort of RGCs in neonatal animals. In neonatal rat visual cortex, extracellular recordings show bursts of spiking activity and slow spreading oscillations that are sensitive to the presence of retinal input in the neocortex *in vivo*[17,18], but patterned spontaneous activity in and among different visual areas associated with retinal waves has never been observed. We sought to establish whether travelling waves of spontaneous retinal activity occur in neonatal mice *in vivo* and to determine whether spontaneous retinal waves convey spatiotemporal patterns suitable for the activity-dependent refinement of visual maps throughout the nervous system before the onset of vision.

Spontaneous retinal waves occur *in vivo*

The most superficial layer of the mouse superior colliculus, the stratum griseum superficiale (SGS), receives retinotopically mapped terminal input from virtually all RGCs[19] and is completely accessible to imaging during development. We recorded directly from RGCs *in vivo* by injecting a calcium indicator, calcium green-1 dextran (CaGr-Dx), into the retina of neonatal mice and imaging the anterogradely labelled RGC axon arbors in the superior colliculus between 3 and 9 days after birth (postnatal day 3 (P3)–P9) through a cranial window using multiphoton or wide-field fluorescence excitation (Fig. 1). Calcium imaging of RGC terminals revealed spontaneous waves of activity propagating throughout the superior colliculus of neonatal mice (Fig. 1d–f and Supplementary Movie 1). Waves propagated among labelled RGC terminals at all depths examined within the SGS (range: 76–206 μm below the pial surface). Wave-front velocity (median ± median absolute deviation: $44.17 \pm 15.99\ \mu\mathrm{m\,s^{-1}}$, $n = 765$ waves) was 3 to 4 times slower than previously reported wave speeds measured in the mouse retina *in vitro*[20,21], which is consistent with the diameter of the superior colliculus being approximately 3 times smaller than that of the retina. Retinal waves occurred with a mean frequency (0.68 ± 0.08 waves per min; $n = 19$ hemispheres) that was sensitive to age (Supplementary Fig. 1) and similar to the frequency observed *in vitro*[20,21]. Surprisingly, waves *in vivo* propagated over a much greater proportion of the superior colliculus area than was expected based on previous *in vitro* studies (median wave size ± median absolute deviation: $0.52 \pm 0.30\ \mathrm{mm}^2$, corresponding to ~30% of superior colliculus retinotopic area, compared with an average wave size of ~1% of retina area *in vitro*[22]) (Supplementary Fig. 2). Waves were completely abolished after application of the Na^+ channel antagonist tetrodotoxin (TTX) into the

---

[1] Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

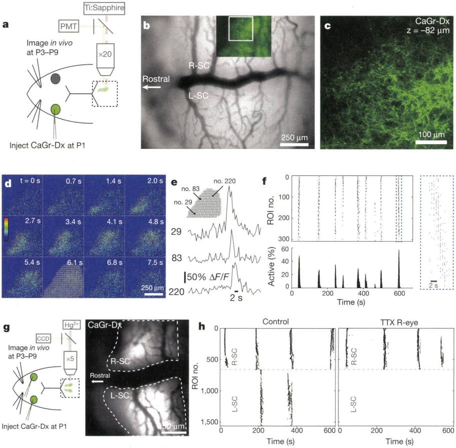


Figure 1 | Spontaneous waves of activity in retinal ganglion cell arbors *in vivo*. **a**, Calcium-dye labelling and two-photon imaging. **b**, Superior colliculus (SC) craniotomy from a P5 mouse overlaid with the corresponding two-photon excitation image of CaGr-Dx-labelled RGC axon arbors. **c**, Higher-magnification image of CaGr-Dx-labelled RGC axons from inset in **b**, 82 μm below pial surface. **d**, Montage ($\Delta F/F$) of a spontaneous wave recorded in RGC axon arbors (from the same labelled field shown in **b**). **e**, Calcium transients from RGC axon arbors during the wave shown in **d**. **f**, Raster plot and activity histogram from a 677-s recording. Points indicate calcium-transient onsets from individual ROIs. Same wave from **d** shown on right at expanded time scale. **g**, Dye labelling and wide-field CCD imaging for bilateral recordings. **h**, Raster plots from animal shown in **g**. Intraocular TTX application blocks waves in the contralateral hemisphere. L-SC, left SC; PMT, photomultiplier tube; R-eye, right eye; R-SC, right SC; Ti:Sapphire, Ti:Sapphire laser.

contralateral eye (Fig. 1g, h and Supplementary Fig. 3) ($n = 2$ animals; region of interest (ROI) active fraction: control, 0.78 ± 0.07; TTX, 0.00 ± 0.00), but not the ipsilateral eye ($n = 2$ animals, ROI active fraction: control, 0.89 ± 0.09; TTX, 0.999 ± 0.001).

Retinal waves propagate to collicular neurons

Most models for activity-dependent neural-circuit development assume that patterned afferent activity drives target-neuron spiking in order for Hebbian plasticity rules to strengthen synapses between co-active neurons. Previous work carried out in an intact mouse retinogeniculate preparation *ex vivo* showed that spontaneous retinal activity can drive dLGN neurons above spike threshold[23] and *in vivo* multi-electrode recordings in the ferret showed that dLGN neurons exhibit correlated episodic bursting that was sensitive to ablation of retinal input[24] just before eye opening. However, it remains unknown whether spontaneous retinal waves *in vivo* can actually drive wave-like network activity in RGC neural targets, a prerequisite for Hebbian refinement at developing synapses. We performed cellular-level calcium imaging in the SGS using bulk loading of the calcium indicator OGB1-AM (Fig. 2a, b) to determine whether retinal waves trigger travelling waves in retinorecipient neurons in the superior colliculus. Bolus-indicator loading in the superficial layer of the superior colliculus labelled >1,000 neurons within a 550 × 550-μm field of view (Fig. 2b, c). We first imaged at low power over the network of cells and neuropil in the same way as for CaGr-Dx imaging (Fig. 1). This revealed propagating calcium waves travelling within the retinorecipient layer of the superior colliculus (range: 34–172 μm below the pial surface) (Fig. 2d–f and Supplementary Movie 2). These waves occurred at the same frequency (0.69 ± 0.17 waves per min, $n = 11$ hemispheres) as the retinal waves recorded presynaptically in RGC axons ($P = 0.9497$, t-test) (Supplementary Fig. 1) and activated a similar percentage of the labelled area in the superior colliculus per recording (CaGr-Dx ROI active fraction: 0.79 ± 0.06, $n = 19$ hemispheres; OGB1-AM ROI active fraction: 0.74 ± 0.08, $n = 11$ hemispheres; $P = 0.6228$, *t*-test). Waves propagated among OGB1-AM-labelled cells in the superior colliculus with a median wave-front velocity of 36.11 μm s^{-1} (median absolute deviation: 13.00 μm s^{-1}, $n = 189$ waves), which was similar to the wave speeds observed with presynaptic recordings (CaGr-Dx mean: 42.93 ± 1.73 μm s^{-1}, OGB1-AM mean: 43.04 ± 6.68 μm s^{-1}, $P = 0.9881$, *t*-test). Together, these results indicate that the spatiotemporal properties of presynaptic retinal waves match those of postsynaptic retinal waves measured in the superior colliculus.

We next performed high spatial resolution two-photon imaging on active regions observed at low magnification. We found that calcium waves propagated among superior colliculus neurons (Fig. 2g–i) (fraction of cells active per wave, 0.48 ± 0.07; fraction of cells active in a wave during a 10-min recording, 0.87 ± 0.08; $n = 787$ cells, 13 waves, 3 recordings), and that the activity of nearby cells was more temporally correlated than distant cells (Fig. 2i, j and Supplementary Fig. 5). The travelling waves among cells in the retinal-input layer of the superior colliculus were abolished when spiking activity was blocked in the contralateral eye (Fig. 2h and Supplementary Fig. 3)

($n = 6$ hemispheres; ROI active fraction: control, 0.80 ± 0.08; TTX, 0.004 ± 0.003) but not the ipsilateral eye ($n = 2$ animals, ROI active fraction: control, 1.00 ± 0.00; TTX, 0.998 ± 0.002). These experiments show that retinal waves drive wave-like activation patterns in retino-recipient target neurons, indicating that spontaneous activity in the retina provides a template pattern that is matched in higher-order circuits in the visual system.

Binocular coordination of retinal waves

Retinal waves are thought to emerge from recurrent excitatory connections in networks of spontaneously active amacrine cells and ganglion cells[25], and are presumed to originate randomly throughout the retina. We examined the initiation sites of retinal waves *in vivo* by analysing the calcium event response frequencies for all active ROIs at the onset of waves. Interestingly, we discovered that instead of a uniform distribution of wave-initiation sites, waves preferentially nucleated in the rostral–medial superior colliculus (Fig. 3a) ($n = 995$ waves). Consistent with this preferential site of wave generation, waves propagated with a marked directional bias towards the caudal–lateral superior colliculus, equivalent to a directional preference for the dorsal–nasal pole of the retina (Fig. 3b) ($n = 995$ waves, Rayleigh test of Uniformity, $P = 0$). This directional bias was the same for both presynaptic and postsynaptic recordings (Supplementary Fig. 6), and was similar to but stronger than the wave-direction bias reported in a recent *in vitro* study[21]. The overall preference for waves to propagate towards the dorsal–nasal retina is aligned with but directly opposite the direction preference of RGCs and visual cortical neurons at the time of eye opening[26,27], which corresponds to movement towards the ventral–temporal pole of the retina. It is possible that the pronounced wave-direction bias we observed mediates the development of aspects of this directionally sensitive visual circuit. The notable preference for wave initiation in the rostral–medial superior colliculus (Fig. 3a) indicates a corresponding site of wave generation in the ventral–temporal retina, which represents the binocular portion of the visual field in mice. This suggests that retinal waves may have an enhanced role in mediating development of binocular aspects of maps in the visual system, such as the formation of eye-specific connectivity and matched orientation selectivity between the eyes.

As retinal waves arise spontaneously, it is commonly assumed that waves in each eye are autonomous and drive their corresponding visual circuits independently, giving rise to a completely asynchronous activation pattern that could be used for segregation of eye-specific projections[6,28]. As unilateral TTX blockade or enucleation abolished presynaptic and postsynaptic retinal waves (Supplementary Fig. 3), we anticipated that the retinal waves would occur independently in the two superior colliculus hemispheres. Unexpectedly, simultaneous

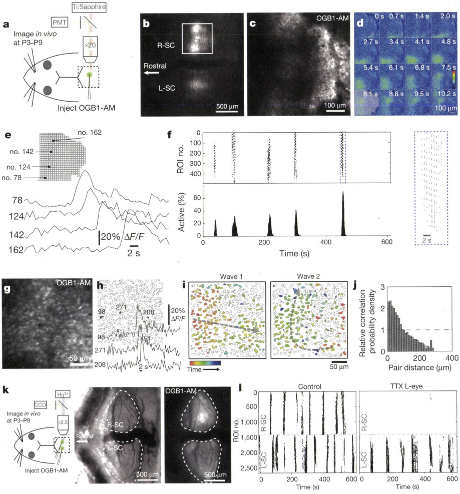


Figure 2 | Spontaneous waves of correlated activity among superior-colliculus neurons *in vivo*. **a**, Experimental overview. **b**, OGB1-AM bulk loading in the SC at P4. **c**, Two-photon excitation image of OGB1-AM-labelled cells from inset in **b**, Image from 56 μm below pial surface. **d**, Montage ($\Delta F/F$) of a spontaneous wave recorded in the SC (from the same labelled field shown in **c**). **e**, Calcium transients in local ROIs during the wave shown in **d**. **f**, Raster plot and activity histogram during a 610-s recording (from the same experiment as **d** and **e**). **g**, Calcium imaging in single SC cells loaded with OGB1-AM. **h**, Single-cell ROIs and calcium transients during a wave. **i**, Plot of neurons active during two sequential waves. **j**, Pairwise cell correlations as a function of pair distance. Dashed line, the chance distribution. **k**, Bilateral dye labelling and example wide-field CCD image. **l**, Raster plots from a bilateral SC recording. Intraocular TTX application blocks waves in the contralateral hemisphere.

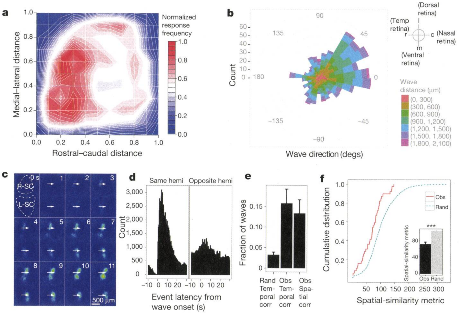


Figure 3 | Retinal waves originate in the ventrotemporal retina and propagate bilaterally. **a**, Contour map of normalized ROI response frequencies in the SC within 2 s of wave onset. **b**, Wind-rose histogram of all wave directions in SC. **c**, Montage ($\Delta F/F$) of synchronized retinal waves recorded in OGB1-AM-labelled cells in both hemispheres of the SC. **d**, Peri-event time histograms of calcium-event latencies relative to wave onset for the same or opposite hemispheres. ROI event latencies: same hemisphere (Same hemi), $n = 103{,}354$; opposite hemisphere (Opp hemi), $n = 97{,}280$. **e**, Fraction of presynaptic and postsynaptic waves having significant temporal ($P = 0.00098$, *t*-test; CaGr-Dx: 52/332 waves; OGB1-AM: 20/90 waves) and spatial ($P = 0.004535$, *t*-test; CaGr-Dx: 46/332 waves; OGB1-AM: 16/90 waves) correlations bilaterally. Error bars, mean ± s.e.m. **f**, Cumulative distributions of the spatial-similarity metric across hemispheres for observed wave pairs and random wave pairs. $^{***}P = 6.072\mathrm{e}^{-06}$, *t*-test.

bilateral calcium imaging of the superior colliculus revealed that retinal waves sometimes travelled with matched spatial patterns in both hemispheres at the same time (Fig. 3c–f). These bilaterally coordinated waves were found when recording from presynaptic afferents in the superior colliculus with CaGr-Dx (Supplementary Movie 3) and from the postsynaptic networks within the superior colliculus labelled with OGB1-AM (Fig. 3c and Supplementary Movie 4). Temporally matched waves occurred throughout the superior colliculus and constituted a small (15.8%) but significant subset of all waves ($P = 0.00098$, *t*-test). Synchronous waves in the two hemispheres of the superior colliculus also typically propagated in a highly correlated manner spatially (85.0% of temporally correlated waves were also spatially correlated, accounting for 13.4% overall; $P = 0.004535$, *t*-test; Fig. 3c–e). Retinopetal projections as well as retino-retinal connectivity have been reported in the developing mammalian retina[29,30] (Supplementary Fig. 7). Thus, it is possible that descending synchronized inputs or synaptic interactions between the two retinas may cause a small subset of retinal waves to initiate at matched retinotopic locations bilaterally. Because commissural connections between the hemispheres in superior colliculus and visual cortex are retinotopically coordinated[31,32] and orientation selectivity is matched between the eyes in the binocular visual cortex, even in the absence of visual experience[33], a possible function for synchronous retinal waves is to regulate bilateral matching of visual map connectivity.

Retinal waves propagate to, and within, visual cortex

Visual information from the retina is routed to the primary visual cortex (V1) through the dLGN of the thalamus. Retinogeniculate input to the dLGN in mice originates from collateral fibres that arise from a subset of retinocollicular afferents. Therefore, retinal waves transmitted to the superior colliculus should also be relayed to the visual cortex through the thalamus. To test whether retinal waves induce propagating waves within visual cortex, we carried out simultaneous wide-field calcium imaging of the superior colliculus and ipsilateral visual cortex (Fig. 4a–c). We found that travelling waves in the visual cortex were coincident with retinal waves propagating across the ipsilateral superior colliculus (Fig. 4d, e) (coactive fraction, V1/superior colliculus = 50/81 waves, $n = 4$ hemispheres), and inter-wave intervals were similar in superior colliculus and V1 (Fig. 4f) (P6–P9 superior colliculus, 40.53 ± 2.81 s, $n = 133$ waves; P6–P9 V1, 49.01 ± 4.79 s, $n = 62$ waves; $P = 0.1292$, *t*-test). The retinotopic map in V1 is mirrored and rotated with respect to the retinotopic map in the ipsilateral superior colliculus (Supplementary Fig. 8). Notably, after correcting for the mirrored and rotated maps, the direction of wave propagation in visual cortex was retinotopically matched to the direction of wave travel in the superior colliculus (Fig. 4g, h) ($r = 0.835$, $P = 4.707\mathrm{e}^{-14}$, Pearson's correlation, $n = 50$ waves). These results indicate that retinal waves carry functional information corresponding to retinal organization simultaneously throughout multiple areas of the visual system during development.

The V1 in mice contains a retinotopic map of the contralateral visual hemifield and is adjoined by up to nine extrastriate visual areas homologous to those seen in carnivore and primates[34], each of which contains a duplicate map of retinal topography. To gain a better understanding of the overall contribution of retinal waves to ongoing patterned activity among developing visual cortical areas in mice, we used a Cre/*loxP* approach (*Emx1-Cre:Ai38* mice) to express the genetically encoded calcium sensor GCaMP3 in all excitatory cortical neurons[35]. Wide-field fluorescent GCaMP3 signals from neocortical neurons were imaged directly through the skull simultaneously with OGB1-AM signals through a cranial window over the superior colliculus (Fig. 5a, b). Retinal waves in the superior colliculus propagated concurrently with waves in the V1 and separate, secondary wave fronts were initiated in extrastriate cortical areas (Fig. 5c and Supplementary Movie 5). We created maps of retinal topography in the visual cortex and midbrain by colorizing and merging $\Delta F/F$ image frames based on the locations of superior colliculus retinal wave fronts within 10-min recordings (Fig. 5d). This functional map of wave-based retinotopy revealed a large primary map in visual cortex that was mirrored and rotated with respect to the superior colliculus map

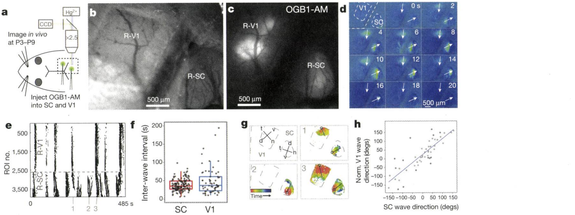


Figure 4 | Retinal waves propagate simultaneously in the SC and visual cortex. a, Experimental overview. **b**, Right visual cortex (R-V1) and SC (R-SC) craniotomies from a P9 mouse. **c**, OGB1-AM bulk loading in V1 and SC. **d**, Montage ($\Delta F/F$) of spontaneous waves recorded in SC and V1 simultaneously. **e**, Raster plot from the recording in **d**. **f**, Box-plots show interval between successive wave onsets in SC and V1 between P6 and P9. Box edges, first and third quartiles. Horizontal line, median. Whiskers, range. **g**, Activity maps showing the direction of wave propagation for 3 waves in V1 and SC indicated in **e**. Top-left panel shows the approximate retinotopy of mouse V1 and SC. d, dorsal; n, nasal; t, temporal; v, ventral. **h**, Plot of SC retinal-wave directions versus V1-wave directions normalized to SC coordinates, with linear regression (blue line) and 95% confidence intervals (grey shading).

(Fig. 5d) and that was remarkably consistent with the known visual-map organization of adult mouse V1 compared with superior colliculus[34,36]. Although the frequency of calcium events was higher in V1 and extrastriate cortex relative to superior colliculus (Fig. 5e) (median ± median absolute deviation, superior colliculus: 0.010 ± 0.003 Hz; V1: 0.013 ± 0.003 Hz; extrastriate: 0.017 ± 0.003 Hz), most of the activity in V1 occurred within seconds of a retinal wave (Fig. 5f). Retinal waves also triggered activity in secondary visual cortical areas, but much of the ongoing cortical activity in extrastriate areas seemed to be independent of retinal waves (Fig. 5e–g), indicating a strong role for intra- or subcortical inputs separate from the primary visual pathway in driving ongoing activity within developing extrastriate cortex. Ablation of contralateral retinal input dramatically diminished spontaneous activity in the superior colliculus and V1, but in surrounding extrastriate visual areas the frequency of ongoing activity was unaffected although the spatial distribution of this activity was shifted noticeably (Fig. 5e, g) (median ± median absolute deviation, superior colliculus: 0 ± 0 Hz; V1: 0.002 ± 0.002 Hz; extrastriate: 0.017 ± 0.005 Hz and Supplementary Fig. 9). These results indicate that retinal waves provide a primary source of patterned activity for V1 during early development and further modulate activity in extrastriate visual areas. Retinal waves convey patterned information suitable for driving activity-dependent development of neocortical circuits before the onset of vision and may thus prompt coordinated alignment and refinement of topographic maps across the visual system.

Retinal waves are abolished by epibatidine

Classic studies show that pharmacologic manipulation of retinal waves *in vivo* with epibatidine, a high-affinity agonist for nicotinic acetylcholine receptors, disrupts visual-map development in ferret[9,10,13,37] and mouse[38–40]. These studies relied on *in vitro* evidence that epibatidine blocks retinal waves[9,41], but the effects of epibatidine on wave generation *in vivo* and the link between retinal waves and visual-map development remain uncertain. We generated *Rx-Cre:Ai38* mice, which express

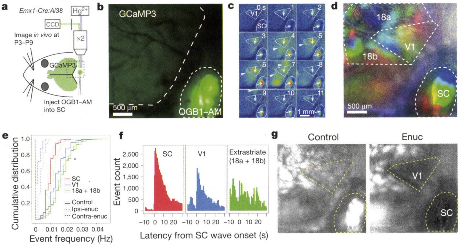


Figure 5 | Retinal-wave-driven activity in V1 and extrastriate visual areas. a, Experimental overview. **b**, Image of transcranial cortical GCaMP3 expression and OGB1-AM bulk loading in SC of a P6 *Emx1-Cre:Ai38* mouse. **c**, Montage ($\Delta F/F$) of a retinal wave recorded in SC and V1 simultaneously from same recording as in **b**. Arrows, direction and distance of wave travel. Arrowheads, activations in V2. **d**, Topographic maps colourized by SC retinal-wave-front position reveal putative boundaries of V1 and secondary visual areas (areas 18a and 18b) from the same animal as in **c**. **e**, Cumulative distributions of calcium-event frequencies before (control) and after ablation (enucleation) of contralateral (Contra-enuc) or ipsilateral (Ipsi-enuc) eye. SC, $n = 2{,}862$ ROIs; V1, $n = 3{,}490$ ROIs; extrastriate, 7,717 ROIs. **f**, Peri-event time histograms of calcium-event latencies relative to SC-wave onset. SC, $n = 52{,}284$ events; V1, $n = 38{,}244$ events; extrastriate, $n = 49{,}363$ events. **g**, Maximum $\Delta F/F$ intensity projections from 10-min recordings before and after contralateral eye enucleation (from the same recording as in **d**).

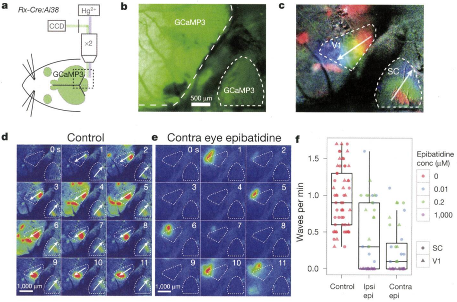


Figure 6 | Retinal waves depend on cholinergic neurotransmission. **a**, Experimental overview. **b**, Image of transcranial cortical GCaMP3 expression and RGC-axon GCaMP3 expression in SC in a P4 *Rx-Cre:Ai38* mouse. **c**, Topographic maps colourized by SC retinal-wave-front position for a single retinal wave simultaneously propagating in SC and V1. **d**, Montage ($\Delta F/F$) of a single retinal wave from same recording as in **c**. **e**, Montage ($\Delta F/F$) showing typical activity patterns in visual cortex after contralateral injection of 1 mM epibatidine (from the same experiment as **d**). **f**, Box-plots showing retinal-wave frequency in 10-min recordings before (control) and after epibatidine injection into contralateral (contra epi; $P = 2.2\mathrm{e}^{-12}$ versus control, pairwise-*t*-test) or ipsilateral (Ipsi epi; $P = 7.3\mathrm{e}^{-08}$ versis control) eye in 6 mice between P2 and P5.

the genetically encoded calcium reporter, GCaMP3, in both RGCs and forebrain neurons to perform simultaneous recordings of retinal waves in RGC axons and the visual cortex. Retinal waves propagated simultaneously in superior colliculus and V1 (Fig. 6c, d) and were completely abolished after monocular injection into the contralateral eye of 1,000 µM epibatidine (Fig. 6e, f) (control, $n = 58$ movies, 0.96 ± 0.05 waves per min; contra epi, $n = 14$ movies, 0 waves per min; $P = 8.2\mathrm{e}^{-14}$, pairwise-*t*-test) (Supplementary Movie 6), the concentration typically used *in vivo* to examine the role of retinal waves in visual-map development[13,37–40]. Notably, retinal waves were also abolished in superior colliculus and V1 ipsilateral to the injected eye (Fig. 6f) ($n = 14$ movies, 0 waves per min; $P = 8.2\mathrm{e}^{-14}$), suggesting that at 1,000 µM epibatidine could mediate some of its effects systemically. Consistent with a concentration-dependent systemic effect, monocular injection of 0.01 µM or 0.2 µM epibatidine significantly reduced wave frequency in the contralateral (0.01 µM, 0.36 ± 0.12 waves per min, $n = 9$ movies, $P = 0.00012$; 0.2 µM, 0.45 ± 0.10 waves per min, $n = 12$ movies, $P = 0.00028$, pairwise-*t*-test) but not the ipsilateral hemisphere (0.01 µM, 0.60 ± 0.18 waves per min, $n = 9$ movies, $P = 0.06280$; 0.2 µM, 0.82 ± 0.08 waves per min, $n = 12$ movies, $P = 0.89135$). These results are consistent with long-standing evidence based on *in vitro* recordings[8] that cholinergic neurotransmission has an essential role in generating retinal waves, and that *in vivo* manipulations with epibatidine disrupt visual-map development by abolishing retinal waves throughout the visual system.

Discussion

There is substantial *in vitro* evidence for spontaneous waves of neuronal activity in the prenatal or postnatal retina of a number of vertebrate species, including rodent, ferret and monkey[42,43]. Slow activity transients or bursting activity measured with electroencephalography (EEG) or extracellular electrodes in the dLGN and visual cortex of rat or ferret *in vivo* before eye opening are modulated by retinal input and therefore believed to be driven by retinal waves[18,24,44]. These slow activity transients are homologous to the slow activity EEG transients recorded in preterm human fetal occipital cortex[45,46]. As these self-organized patterns of spontaneous activity are present before the onset of patterned vision, it has long been thought that retinal waves may help organize aspects of visual circuit function that are established before visual experience, such as maps for orientation[33,47], direction[27] and ocular dominance[33,48].

Our recordings in mouse pups aged P3–P9 indicate that spontaneous retinal waves are present for at least a week of development *in vivo* and exhibit a pattern of activity appropriate for communicating retinal organization to circuits throughout the visual system. Imaging of RGC afferents or superior colliculus neurons together with simultaneous imaging in V1 shows that retinal waves generate matched activity patterns in the midbrain and cortex, supporting the hypothesis that emergent retinal activity mediates the development of linked visual circuitry within and across multiple neocortical and subcortical brain regions[49].

Retinal waves *in vivo* exhibited a number of properties that could not have been predicted based on previous *in vitro* recordings, including a preferential site of generation in the ventral–temporal retina, a biased direction of wave travel towards the dorsal–nasal retina, and a differential role in driving primary versus secondary visual cortical activities (see also Supplementary Discussion). Unexpectedly, we also found that a subset of retinal waves exhibited coordinated patterns across both hemispheres. This suggests that bilateral retinal waves could help establish inter-hemispheric connections, a role not previously ascribed to waves. Moreover, the remarkable mirror-symmetric pattern of ocular-dominance columns between hemispheres observed in some monkeys[50], which are difficult to explain based on the presumed random pattern of waves between the eyes, may be understood better in the context of the clearly non-random nature of waves across the two hemispheres found *in vivo*.

A causal link between retinal waves and visual-map development has been difficult to establish definitively, partly owing to the lack of evidence for wave generation and propagation *in vivo*. Our experiments show that pharmacological manipulations that were previously shown to profoundly disrupt visual-map development[9,10,13,37–39] also block wave generation *in vivo*, strengthening the causal link between wave activity and map development.

Given the remarkable fidelity of retinal waves during the period before eye opening in mice *in vivo* that we report here, together with previous work showing that spontaneous waves within macaque retina are present *in vitro* before birth[43], it seems likely that the visual system experiences patterned activation by retinal waves for a substantial gestational period during human fetal development that may be crucial for shaping the functional maturation of neural circuits before the onset of sensory experience.

METHODS SUMMARY

Anaesthetized C57BL/6 mice between P1 and P3 were injected with calcium green-1 dextran (CaGr-Dx) into each retina and returned to the dam for at least 2 days before preparation for *in vivo* imaging of RGC axons. On the day of calcium imaging, a craniotomy was prepared over the superior colliculus under anaesthesia, and a cover-glass was placed over the superior colliculus in warm agarose-buffered saline, pH 7.4.

For postsynaptic calcium imaging, craniotomies were prepared in anaesthetized mice aged P3 to P9 over the superior colliculus and visual cortex. The calcium-sensitive dye oregon green BAPTA-1-AM (OGB1-AM) was injected into the superior colliculus and visual cortex and then a cover-glass was placed over the craniotomies. *In vivo* calcium imaging was carried out in head-fixed, un-anaesthetized mouse pups using two-photon excitation through a ×20, 1.0-NA (numerical aperture) objective lens or using wide-field epifluorescence imaging with low-magnification objectives (×2–5) and a charge-coupled device (CCD) camera when the animals awoke during or after a 1-h recovery period from general anaesthesia. Calcium transients and waves were detected and analysed using custom routines implemented in MATLAB and R.

Full Methods and any associated references are available in the online version of the paper.

Received 27 April; accepted 20 August 2012.

1. Buffelli, M. *et al.* Genetic evidence that relative synaptic efficacy biases the outcome of synaptic competition. *Nature* **424,** 430–434 (2003).
2. Marder, E. & Rehm, K. J. Development of central pattern generating circuits. *Curr. Opin. Neurobiol.* **15,** 86–93 (2005).
3. Petersson, P., Waldenström, A., Fåhraeus, C. & Schouenborg, J. Spontaneous muscle twitches during sleep guide spinal self-organization. *Nature* **424,** 72–75 (2003).
4. Sanes, J. R. & Lichtman, J. W. Development of the vertebrate neuromuscular junction. *Annu. Rev. Neurosci.* **22,** 389–442 (1999).
5. Katz, L. C. & Shatz, C. J. Synaptic activity and the construction of cortical circuits. *Science* **274,** 1133–1138 (1996).
6. Meister, M., Wong, R. O., Baylor, D. A. & Shatz, C. J. Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* **252,** 939–943 (1991).
7. Wong, R. O., Meister, M. & Shatz, C. J. Transient period of correlated bursting activity during development of the mammalian retina. *Neuron* **11,** 923–938 (1993).
8. Feller, M. B., Wellis, D. P., Stellwagen, D., Werblin, F. S. & Shatz, C. J. Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. *Science* **272,** 1182–1187 (1996).
9. Penn, A. A., Riquelme, P. A., Feller, M. B. & Shatz, C. J. Competition in retinogeniculate patterning driven by spontaneous activity. *Science* **279,** 2108–2112 (1998).
10. Stellwagen, D. & Shatz, C. J. An instructive role for retinal waves in the development of retinogeniculate connectivity. *Neuron* **33,** 357–367 (2002).
11. McLaughlin, T., Torborg, C. L., Feller, M. B. & O'Leary, D. D. M. Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. *Neuron* **40,** 1147–1160 (2003).
12. Chandrasekaran, A. R., Plas, D. T., Gonzalez, E. & Crair, M. C. Evidence for an instructive role of retinal activity in retinotopic map refinement in the superior colliculus of the mouse. *J. Neurosci.* **25,** 6929–6938 (2005).
13. Huberman, A. D., Speer, C. M. & Chapman, B. Spontaneous retinal activity mediates development of ocular dominance columns and binocular receptive fields in v1. *Neuron* **52,** 247–254 (2006).
14. Chalupa, L. M. Retinal waves are unlikely to instruct the formation of eye-specific retinogeniculate projections. *Neural Dev.* **4,** 25 (2009).
15. Feller, M. B. Retinal waves are likely to instruct the formation of eye-specific retinogeniculate projections. *Neural Dev.* **4,** 24 (2009).
16. Maffei, L. & Galli-Resta, L. Correlation in the discharges of neighboring rat retinal ganglion cells during prenatal life. *Proc. Natl Acad. Sci. USA* **87,** 2861–2864 (1990).
17. Hanganu, I. L., Ben-Ari, Y. & Khazipov, R. Retinal waves trigger spindle bursts in the neonatal rat visual cortex. *J. Neurosci.* **26,** 6728–6736 (2006).
18. Colonnese, M. T. & Khazipov, R. "Slow activity transients" in infant rat visual cortex: a spreading synchronous oscillation patterned by retinal waves. *J. Neurosci.* **30,** 4325–4337 (2010).
19. Hofbauer, A. & Dräger, U. C. Depth segregation of retinal ganglion cells projecting to mouse superior colliculus. *J. Comp. Neurol.* **234,** 465–474 (1985).
20. Bansal, A. *et al.* Mice lacking specific nicotinic acetylcholine receptor subunits exhibit dramatically altered spontaneous activity patterns and reveal a limited role for retinal waves in forming ON and OFF circuits in the inner retina. *J. Neurosci.* **20,** 7672–7681 (2000).
21. Stafford, B. K., Sher, A., Litke, A. M. & Feldheim, D. A. Spatial-temporal patterns of retinal waves underlying activity-dependent refinement of retinofugal projections. *Neuron* **64,** 200–212 (2009).
22. Ford, K. J., Félix, A. L. & Feller, M. B. Cellular mechanisms underlying spatiotemporal features of cholinergic retinal waves. *J. Neurosci.* **32,** 850–863 (2012).
23. Mooney, R., Penn, A. A., Gallego, R. & Shatz, C. J. Thalamic relay of spontaneous retinal activity prior to vision. *Neuron* **17,** 863–874 (1996).
24. Weliky, M. & Katz, L. C. Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus *in vivo*. *Science* **285,** 599–604 (1999).
25. Blankenship, A. G. & Feller, M. B. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Rev. Neurosci.* **11,** 18–29 (2010).
26. Elstrott, J. *et al.* Direction selectivity in the retina is established independent of visual experience and cholinergic retinal waves. *Neuron* **58,** 499–506 (2008).
27. Rochefort, N. L. *et al.* Development of direction selectivity in mouse cortical neurons. *Neuron* **71,** 425–432 (2011).
28. Zhang, J., Ackman, J. B., Xu, H.-P. & Crair, M. C. Visual map development depends on the temporal pattern of binocular activity in mice. *Nature Neurosci.* **15,** 298–307 (2011).
29. Gastinger, M. J., Tian, N., Horvath, T. & Marshak, D. W. Retinopetal axons in mammals: emphasis on histamine and serotonin. *Curr. Eye Res.* **31,** 655–667 (2006).
30. Müller, M. & Holländer, H. A small population of retinal ganglion cells projecting to the retina of the other eye. An experimental study in the rat and the rabbit. *Exp. Brain Res.* **71,** 611–617 (1988).
31. Takahashi, M., Sugiuchi, Y. & Shinoda, Y. Commissural mirror-symmetric excitation and reciprocal inhibition between the two superior colliculi and their roles in vertical and horizontal eye movements. *J. Neurophysiol.* **98,** 2664–2682 (2007).
32. Lewis, J. W. & Olavarria, J. F. Two rules for callosal connectivity in striate cortex of the rat. *J. Comp. Neurol.* **361,** 119–137 (1995).
33. Crair, M. C., Gillespie, D. C. & Stryker, M. P. The role of visual experience in the development of columns in cat visual cortex. *Science* **279,** 566–570 (1998).
34. Wang, Q. & Burkhalter, A. Area map of mouse visual cortex. *J. Comp. Neurol.* **502,** 339–357 (2007).
35. Zariwala, H. A. *et al.* A Cre-dependent GCaMP3 reporter mouse for neuronal imaging *in vivo*. *J. Neurosci.* **32,** 3131–3141 (2012).
36. Dräger, U. C. & Hubel, D. H. Topography of visual and somatosensory projections to mouse superior colliculus. *J. Neurophysiol.* **39,** 91–101 (1976).
37. Huberman, A. D., Stellwagen, D. & Chapman, B. Decoupling eye-specific segregation from lamination in the lateral geniculate nucleus. *J. Neurosci.* **22,** 9419–9429 (2002).
38. Rossi, F. M. *et al.* Requirement of the nicotinic acetylcholine receptor β2 subunit for the anatomical and functional development of the visual system. *Proc. Natl Acad. Sci. USA* **98,** 6453–6458 (2001).
39. Pfeiffenberger, C. *et al.* Ephrin-As and neural activity are required for eye-specific patterning during retinogeniculate mapping. *Nature Neurosci.* **8,** 1022–1027 (2005).
40. Rebsam, A., Petros, T. J. & Mason, C. A. Switching retinogeniculate axon laterality leads to normal targeting but abnormal eye-specific segregation that is activity dependent. *J. Neurosci.* **29,** 14855–14863 (2009).
41. Sun, C., Speer, C. M., Wang, G.-Y., Chapman, B. & Chalupa, L. M. Epibatidine application *in vitro* blocks retinal waves without silencing all retinal ganglion cell action potentials in developing retina of the mouse and ferret. *J. Neurophysiol.* **100,** 3253–3263 (2008).
42. Wong, R. O. Retinal waves and visual system development. *Annu. Rev. Neurosci.* **22,** 29–47 (1999).
43. Warland, D. K., Huberman, A. D. & Chalupa, L. M. Dynamics of spontaneous activity in the fetal macaque retina during development of retinogeniculate pathways. *J. Neurosci.* **26,** 5190–5197 (2006).
44. Chiu, C. & Weliky, M. Spontaneous activity in developing ferret visual cortex *in vivo*. *J. Neurosci.* **21,** 8906–8914 (2001).
45. Vanhatalo, S. *et al.* Slow endogenous activity transients and developmental expression of K^+-Cl^- cotransporter 2 in the immature human cortex. *Eur. J. Neurosci.* **22,** 2799–2804 (2005).
46. Tolonen, M., Palva, J. M., Andersson, S. & Vanhatalo, S. Development of the spontaneous activity transients and ongoing cortical activity in human preterm babies. *Neuroscience* **145,** 997–1006 (2007).
47. Wiesel, T. N. & Hubel, D. H. Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* **158,** 307–318 (1974).
48. Rakic, P. Prenatal genesis of connections subserving ocular dominance in the rhesus monkey. *Nature* **261,** 467–471 (1976).
49. Triplett, J. W. *et al.* Retinal input instructs alignment of visual topographic maps. *Cell* **139,** 175–185 (2009).
50. Adams, D. L. & Horton, J. C. Capricious expression of cortical columns in the primate brain. *Nature Neurosci.* **6,** 113–114 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Sachdev and D. McCormick for *Emx1-Cre:Ai38* mice, C. Chen for helpful advice on ganglion cell loading with calcium indicators and Y. Zhang for technical support. We would like to thank M. Colonnese and members of the Crair laboratory for valuable comments on the manuscript. This work was supported by US National Institutes of Health (NIH) grants P30 EY000785 and R01 EY015788 (to M.C.C.), T32 NS007224 (to J.A.), and T15 LM070506 and T32 EY017353 (to T.B.). This work was also supported by the family of William Ziegler III.

Author Contributions J.B.A. and M.C.C. designed the experiments. J.B.A. carried out *in vivo* ganglion-cell-axon, collicular-neuron and visual-cortex imaging experiments and analysed the recordings. T.J.B. carried out intra-ocular ganglion-cell labelling and *in vivo* ganglion-cell-axon imaging experiments and analysed recordings. J.B.A. implemented analysis routines and analysed the data. J.B.A. and M.C.C. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.C. (michael.crair@yale.edu).

doi:10.1038/nature11526

A neural circuit for spatial summation in visual cortex

Hillel Adesnik[1]†, William Bruns[1], Hiroki Taniguchi[2], Z. Josh Huang[2] & Massimo Scanziani[1]

The response of cortical neurons to a sensory stimulus is modulated by the context. In the visual cortex, for example, stimulation of a pyramidal cell's receptive-field surround can attenuate the cell's response to a stimulus in the centre of its receptive field, a phenomenon called surround suppression. Whether cortical circuits contribute to surround suppression or whether the phenomenon is entirely relayed from earlier stages of visual processing is debated. Here we show that, in contrast to pyramidal cells, the response of somatostatin-expressing inhibitory neurons (SOMs) in the superficial layers of the mouse visual cortex increases with stimulation of the receptive-field surround. This difference results from the preferential excitation of SOMs by horizontal cortical axons. By perturbing the activity of SOMs, we show that these neurons contribute to pyramidal cells' surround suppression. These results establish a cortical circuit for surround suppression and attribute a particular function to a genetically defined type of inhibitory neuron.

Visual stimuli located outside of the classical receptive field of a neuron in visual cortex are unable to elicit spiking, but they may modulate the neuron's response to stimuli located in its receptive field[1–3]. Surround suppression, a basic operation in visual processing, is a classical example of this type of modulation[4–9] and can be easily observed when monitoring the firing of a neuron to a stimulus of increasing size centred on its receptive field (that is, the size tuning of the neuron): the neuron's initial increase in firing is followed by a decrease in firing as the stimulus becomes progressively larger. This form of suppression has been suggested to contribute to a number of perceptual effects like pop-out, curvature detection and orientation discrimination[8,10–12]. Importantly surround suppression is not only observed in the cortex but is already present at earlier stages along the visual hierarchy, namely in the retina[13–15] and the thalamus[14,16–18]. Thus, although it is likely that at least part of the suppressive surround observed in the cortex is relayed from earlier stages of visual processing[19], some experimental observations and theoretical models suggest that the cortex is itself capable of contributing to surround suppression[20–22]. Here we reveal the identity and describe the mechanism of a cortical circuit that directly contributes to surround suppression in the superficial layers of the primary visual cortex.

Neuron-type-specific spatial summation

We determined the tuning to the size of a visual stimulus for neurons in the superficial layers of the primary visual cortex (V1; depth ~100–350 μm, corresponding approximately to layers 2 and 3 (layer 2/3)) of mice. Experiments were performed in awake, running animals, as size tuning was affected by anaesthesia (Supplementary Fig. 1). Mice were head fixed but otherwise unrestrained and free to run on a passive circular treadmill. To reduce variability due to differences in behavioural state, all data presented here were recorded during running events (see Supplementary Methods).Visual stimuli were composed of circular patches of drifting gratings at maximal contrast, presented in 6 or 7 different sizes (from 8 to 97 degrees in diameter, Fig. 1a). The size-tuning curve of isolated units[23] ($n = 53$; Supplementary Fig. 2), that is, the neuronal firing rate as a function of stimulus size, peaked at 22 ± 2 degrees (preferred size; ± s.e.m.) and progressively decreased with larger stimuli (Fig. 1a, d), revealing marked surround suppression[24]. The mean firing rates (± s.e.m.) were 0.47 ± 0.11 Hz (at the baseline ($\mathrm{FR_{BL}}$)); 3.0 ± 1.0 Hz in response to the smallest stimulus ($\mathrm{FR_{SS}}$); 3.1 ± 0.3 Hz in response to the preferred stimulus ($\mathrm{FR_{PS}}$); and 1.0 ± 0.2 Hz in response to the largest stimulus ($\mathrm{FR_{LS}}$). The stimulus modulation index (SMI, computed as $(\mathrm{FR_{PS}} - \mathrm{FR_{BL}})/\mathrm{FR_{PS}}$) was 0.87 ± 0.03. The suppression index, that is, the difference between the peak response and the response to the largest stimulus, divided by the baseline subtracted peak response ($(\mathrm{FR_{PS}} - \mathrm{FR_{LS}})/(\mathrm{FR_{PS}} - \mathrm{FR_{BL}})$; Fig. 1a) averaged 0.9 ± 0.1 ($n = 53$; the suppression index was statistically significant in 33 out of 53 units (permutation test); Fig. 1d), indicating substantial suppression in response to large stimuli. Infrequent eye movements occurring during running had little effect on the size-tuning curve (Supplementary Fig. 3).

If cortical circuits contribute to surround suppression they may involve the suppressive action of cortical inhibitory neurons. An inhibitory neuron lacking surround suppression and whose response increases with stimulus size would be a good candidate. In cat visual cortex, for example, fast spiking inhibitory neurons respond with higher firing rates to large stimuli than to small visual stimuli[22]. We carried out targeted loose-patch recordings from inhibitory neurons in layer 2/3 of V1 in awake, running mice, using two-photon laser scanning microscopy[25] (Fig. 1b, c). Parvalbumin-expressing neurons (PVs)[26], a large class of cortical inhibitory neurons, were visualized in layer 2/3 by crossing a PV–Cre mouse line with a tdTomato-reporter line. The size-tuning curve of PVs peaked at 57 ± 8 degrees ($n = 11$) and showed marked surround suppression with larger stimuli (suppression index, 0.46 ± 0.12, $n = 11$; suppression index statistically significant in 6 out of 11 cells, permutation test (Fig. 1b, e); $\mathrm{FR_{BL}}$, 9 ± 2 Hz; $\mathrm{FR_{SS}}$, 27 ± 7 Hz; $\mathrm{FR_{PS}}$, 45 ± 11 Hz; $\mathrm{FR_{LS}}$, 26 ± 8 Hz; SMI, 0.74 ± 0.07; recorded PVs showed their characteristic 'thin' spike shapes (Supplementary Fig. 2), confirming the accuracy of our targeting strategy[23,27]). In contrast to PVs, SOMs, another large class of cortical inhibitory neurons[26] (visualized by crossing a SOM–Cre line with a tdTomato-reporter line[28]), completely lacked surround suppression (suppression index, 0.09 ± 0.06; $n = 8$; suppression index statistically

---

[1]Howard Hughes Medical Institute, Center for Neural Circuits and Behavior, Neurobiology Section and Department of Neuroscience, University of California San Diego, La Jolla, California 92093-0634, USA. [2]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. †Present address: Department of Molecular and Cell Biology and the Helen Wills Neuroscience Institute, University of California, Berkeley 94720, USA.

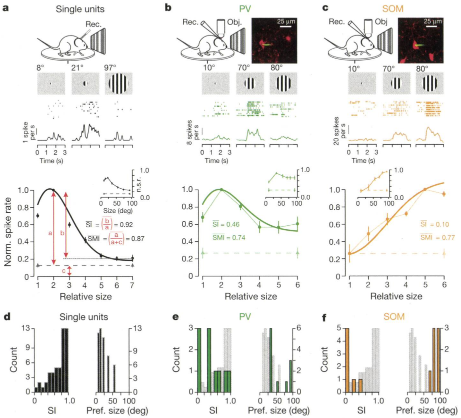


Figure 1 | Differences in spatial summation between three types of neurons in the visual cortex. a, Top, schematic of the experimental setup. Centre, response of example unit to visual stimuli of three different sizes (a raster plot and a peristimulus time histogram (PSTH) are shown). Shaded area behind the raster plot represents the period of stimulus presentation. Bottom, average size-tuning curve of 53 peak-aligned and normalized single units (6 animals, 11 recording sessions). Grey triangles with dashed lines, baseline firing rate. Inset, average of the normalized but not peak-aligned 53 size-tuning curves. SI, suppression index; SMI, stimulus modulation index. **b**, Top left, schematic of the experimental setup. Top right, tdTomato-expressing PV (red) with attached Alexafluor-488-filled recording pipette (green). Center, response of PV to visual stimuli of three different sizes (a raster plot and a PSTH are shown). Bottom, average size-tuning curve ($n = 11$ peak-aligned and normalized size-tuning curves; 3 animals). Inset, average of the normalized but not peak-aligned 11 size-tuning curves. Obj., microscope objective. **c**, As in **b** but for SOMs ($n = 7$; 4 animals). **d–f**, Distribution of SIs (left panels) and of preferred stimulus sizes (right panels) for single units (**d**), PVs (**e**) and SOMs (**f**). The SOM and PV data are superimposed onto the single-unit data (grey, from **e**) for comparisons. All error bars are s.e.m.

significant in 0 out of 8 cells; permutation test; significantly different from PVs; $P < 0.03$, rank sum test; Fig. 1f). The size-tuning curve of these neurons showed a monotonic increase or saturation in firing rate with stimulus size (Fig. 1c, f). Although the smallest stimuli were relatively inefficient in driving SOMs ($\mathrm{FR}_{\mathrm{BL}}$, 7 ± 2 Hz; $\mathrm{FR}_{\mathrm{SS}}$, 5 ± 2 Hz), they robustly responded to large stimuli ($\mathrm{FR}_{\mathrm{PS}}$, 26 ± 2 Hz; preferred size, 86 ± 3 degrees, different from PV, $P < 0.015$, rank sum test; SMI, 0.75 ± 0.05). These data demonstrate that in V1, spatial summation can be very different between genetically distinct types of neurons. Furthermore, these data suggest that SOMs are potential candidates in the generation of cortical surround suppression.

Excitation of SOMs by horizontal axons

In response to the observations above, we examined two fundamental issues: first, what cortical circuits enable SOMs, in contrast to other cortical neurons, to be facilitated rather than suppressed by large stimuli; and second, whether SOMs contribute to size tuning in V1.

The two predominant excitatory inputs to layer 2/3 are vertically ascending axons from layer 4 and horizontally projecting axons from layer 2/3. To examine whether SOMs are equally excited by these two inputs, we recorded from layer 2/3 pyramidal cells, SOMs and PVs in coronal slices of V1 and selectively photo-activated (light ramp of 2 s duration; 480 nm light) layer 4 excitatory cells that conditionally expressed channelrhodopsin 2 (ChR2; Supplementary Fig. 4)[29]. Layer 4 photostimulation generated excitatory charges in SOMs that were only $17 \pm 5\%$ ($\pm$ s.e.m.; $n = 8$) of those in simultaneously recorded pyramidal cells (Fig. 2a); in contrast, excitatory charges generated in PVs were $250 \pm 39\%$ ($n = 8$) of those generated in simultaneously recorded pyramidal cells (Fig. 2b). These results were corroborated through ChR2-assisted circuit mapping[30] (Supplementary Fig. 5). Thus, ascending layer 4 axons provide little excitation to SOMs. Notably, photostimulation of layer 2/3 pyramidal cells (2 s light ramp duration) that selectively expressed ChR2 (see methods[31]) produced substantial excitation of layer 2/3 SOMs (Fig. 2d): the excitation that SOMs received was in fact significantly larger than that received by simultaneously recorded pyramidal cells ($241 \pm 85\%$, $n = 7$, $P < 0.05$; Fig. 2e; we selectively recorded from pyramidal cells that did not express ChR2 in order to avoid contamination with photocurrents[31]). Furthermore, although photostimulation of horizontal layer 2/3 projections was accompanied by strong disynaptic inhibition in pyramidal cells, very little inhibition was recorded in SOMs (Fig. 2e; the ratio of excitation to the sum of excitation and inhibition ($\mathrm{E}/(\mathrm{E}+\mathrm{I})$) was 0.11 ± 0.01 ($n = 10$) in pyramidal cells versus 0.59 ± 0.06 ($n = 11$) in SOMs; $P < 0.05$, Fig. 2f). These results show that although layer 2/3 pyramidal cells and PVs receive substantial excitatory drive from ascending layer 4 axons, the main excitation to layer 2/3 SOMs are horizontal axons of layer 2/3.

Size-dependent excitation of SOMs

To ascertain whether these horizontal axons are indeed responsible for the size-dependent recruitment of SOMs, we took advantage of the retinotopic organization of V1 (ref. 32); we reasoned that because

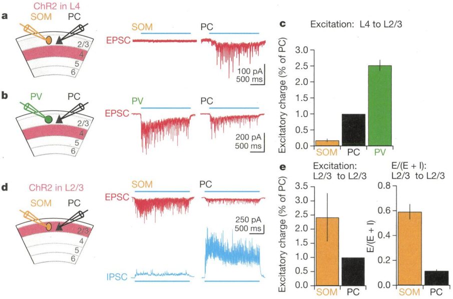


Figure 2 | SOMs are selectively excited by horizontal cortical projections. **a**, Left, schematic of the experimental setup. ChR2 is expressed selectively in layer 4 (L4) excitatory neurons. Recording electrodes in layer 2/3 (L2/3) target a SOM and a pyramidal cell (PC). Right, excitatory currents simultaneously recorded in a SOM and a PC in response to photostimulation of layer 4 with a ramp of blue light (horizontal blue line). EPSC, excitatory postsynaptic current. **b**, Left, schematic of the experimental setup. As in **a** but whole-cell recording electrodes in layer 2/3 target a PV and a PC. Right, excitatory currents simultaneously recorded in a PV and a PC in response to photostimulation of layer 4 as in **a**. **c**, Summary statistics of the excitatory charge (as a fraction of that simultaneously recorded in the PC) recorded in SOMs ($n = 8$) and PVs ($n = 8$); $P < 0.05$. **d**, Left, schematic of the experimental setup. As in **a** except ChR2 is expressed selectively in PCs of layer 2/3. Right, excitatory currents (red, top traces) and inhibitory currents (bottom, blue traces) simultaneously recorded in a SOM and a PC in response to photostimulation of layer 2/3 with a ramp of blue light. IPSC, inhibitory postsynaptic current. **e**, Left, summary statistics of excitatory charge (as a fraction of that simultaneously recorded in the PC) recorded in SOMs compared to that recorded in layer 2/3 PCs ($n = 7$, $P < 0.05$). Right, ratio of excitation to inhibition (expressed as E/E + I) recorded in SOMs and PCs ($n = 10$, $P < 0.05$). All error bars are s.e.m.

progressively larger visual stimuli presented *in vivo* will result in a progressively larger visually activated area in V1, we could approximate this expansion of activity by directly photostimulating progressively larger areas of V1. We carried out loose-patch recordings from SOMs in coronal slices of V1 expressing ChR2 in layer 2/3 pyramidal cells (Fig. 3a). The firing rate of SOMs increased as a function of the size of the light spot (180 to 900 μm), similar to their increase in firing rate *in vivo* with increasing visual stimulus size (Fig. 3a, b). Consistent with the increase in firing rate, the synaptic excitation received by SOMs increased with increasing light spot size (Fig. 3a, b). If SOM dendrites were to span areas similar to the largest light-spot diameter, the progressive increase in firing rate with spot size might simply result from the direct photostimulation of synapses on the dendritic arborization of the recorded SOM. However, this was not the case because even the smallest light spot used (180 μm diameter), generating only approximately 25% of the maximal firing rate, covered more than 95% of the entire SOM dendritic arborization (Fig. 3b, see Supplementary Methods). Thus, the increase in SOM firing rate with spot size results from the recruitment of progressively more distant L2/3 pyramidal cells (Fig. 3b, see methods). Furthermore, cutting horizontal axons with two vertical cuts through layer 2/3 on each side of the recorded SOMs (320 ± 25 μm (± s.e.m.) between cuts, centred on the cell, $n = 10$; note that the distance between the cuts is larger than the horizontal dendritic extent of SOMs) prevented the increase in firing rate with stimuli larger than the distance between the two cuts (Supplementary Fig. 6). Thus, by using horizontal layer 2/3 projections as their main excitatory drive, SOMs are recruited as a function of the activated V1 area; that is, they sum activity in visual space.

Is the size-dependent recruitment of SOMs a mechanism that could contribute to the suppression of pyramidal-cell firing to large stimuli? We recorded from pyramidal cells in coronal slices of V1 conditionally expressing ChR2 in layer 2/3 (again, we selectively recorded from pyramidal cells that did not express ChR2). The firing rate of pyramidal cells was set to approximately 10 Hz by direct-current injection (Fig. 3c). A small light spot centred around the recorded pyramidal cell reduced the firing rate, and this suppression became progressively more pronounced as a larger area of layer 2/3 was activated by increasing the size of the light spot (Fig. 3c). Consistent with the progressive suppression in firing rate, the inhibition received by pyramidal cells increased with increasing light-spot size (Fig. 3c,d). Finally, to establish that the inhibition generated in pyramidal cells after photo-activation of layer 2/3 (ref. 31) was due to the recruitment of SOMs[33] and not any other interneuron type, we optogenetically silenced SOMs (see Methods) while monitoring the inhibition in pyramidal cells during photoactivation of layer 2/3 (Fig. 3e). Photostimulation of layer 2/3 to activate pyramidal cells generated strong firing in SOMs and large inhibitory currents in simultaneously recorded pyramidal cells, consistent with the results reported above (Fig. 3e). Notably, concomitant optogenetic silencing of SOMs (100% reduction of firing; $n = 6$) strongly reduced the inhibitory currents in pyramidal cells (80 ± 4% reduction (± s.e.m.); $n = 8$, $P < 0.05$, Fig. 3e, f).

Thus, the stimulus-size-dependent recruitment of SOMs generates strong inhibition in layer 2/3 pyramidal cells and efficiently suppresses their firing rate (Fig. 3g).

SOMs contribute to surround suppression

These data provide a plausible mechanism by which SOMs could contribute to surround suppression of layer 2/3 pyramidal cells *in vivo*. Furthermore, under anaesthesia, a situation in which surround suppression is compromised (see above and Supplementary Fig. 1), the firing rate of SOMs was reduced tenfold (from 26 ± 2 Hz, $n = 8$ to 2.7 ± 0.4 Hz, $n = 10$), much more than that of single units or PVs (Supplementary Fig. 1). This is consistent with a possible contribution of SOMs to surround suppression. To test directly for the involvement of SOMs in surround suppression, we conditionally expressed the light-sensitive hyperpolarizing opsin archaerhodopsin (Arch)[34] in V1 using viral injection of a flexed Arch vector[35]into SOM-Cre mice (71 ± 2% of cells infected (± s.e.m.), $n = 4$ animals; Fig. 4b and Methods). Illumination of the cortical surface efficiently reduced the

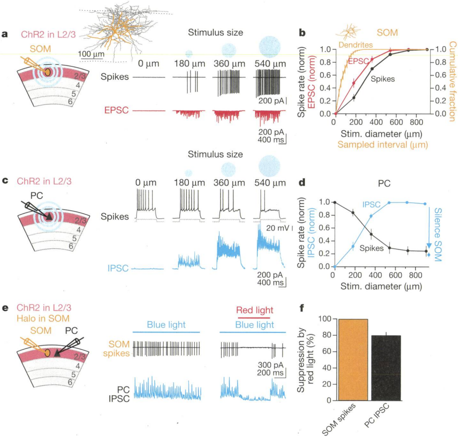


Figure 3 | Suppression of pyramidal cells by SOMs as a function of the activated layer 2/3 area. a, Left, schematic of the experimental setup. ChR2 is expressed in a fraction of PCs in layer 2/3. Inset, anatomical reconstruction of a biocytin-filled layer 2/3 SOM (dendrites, orange; axon, grey; top dotted line, border with layer 1; bottom dotted line, border with layer 4). Right, action potentials (black traces, top) recorded in the cell-attached mode in a SOM in response to light-spot sizes of increasing diameter. Bottom, excitatory currents (red traces) recorded subsequently in the whole-cell voltage-clamp configuration in the same SOM neuron in response to the same photo-stimuli. EPSC, excitatory postsynaptic current. **b**, Summary graph for the spiking (black; $n = 14$) and excitatory charge (red; $n = 6$) of SOMs in response to light spots of five different diameters. Orange, summary statistics of the cumulative fraction of SOM dendritic arbor length within a sampled horizontal interval centred on the SOM cell body ($n = 6$). Inset, dendrites of the SOM illustrated in **a** but scaled to x axis. **c**, Left, schematic of the experimental setup. As in **a** but recording from PC. Right, spiking of PC recorded in current-clamp mode (black traces) in response to depolarizing current steps while layer 2/3 is photo-stimulated with increasingly large blue light spots (top), and inhibitory currents recorded in a PC to the same light stimuli (bottom). IPSC, inhibitory postsynaptic current. **d**, Summary graph of the suppression of firing of PCs (black, $n = 7$) and intracellularly recorded inhibitory charge (blue, $n = 6$) to light spots of five different diameters. Photo-hyperpolarizing SOMs (blue arrow; see **e**) reduce inhibitory charge in PCs. **e**, Schematic of the experimental setup. ChR2 is expressed in a fraction of layer 2/3 PCs and halorhodopsin (halo) is conditionally expressed in SOMs. Recording electrodes target a SOM and a PC. Full-field blue light activates layer 2/3 PCs, whereas red light suppress SOMs. Traces, spikes (black traces, top) recorded in the cell-attached mode in a SOM and inhibitory currents (blue traces, bottom) simultaneously recorded in a voltage-clamped PC in response to blue light photo-stimulation (blue bar) of layer 2/3. Simultaneous illumination with red light (red bar, right panel) to photo-hyperpolarize SOMs abolishes SOM firing and reduces inhibitory currents in the PC (see also blue arrow in **d**). **f**, Summary graph for halorhodopsin-mediated reduction of SOM firing ($n = 6$) and concomitant reduction in inhibitory charge in layer 2/3 PCs ($n = 8$).

visually evoked activity of Arch-expressing layer 2/3 SOMs (80 ± 1% suppression, $n = 4$, $P < 0.05$; Supplementary Fig. 7). To determine the impact of SOMs on size tuning in layer 2/3, we performed extracellular recordings as described above and alternated control trials (visual stimulus only) with trials in which SOMs were photo-hyperpolarized (Fig. 4). Photo-hyperpolarization of SOMs significantly reduced surround suppression of layer 2/3 neurons by 30 ± 10% (± s.e.m.; $n = 28$, $P < 0.00022$, paired signed rank test, Fig. 4c–e; photo-hyperpolarization of SOMs had no significant effect on baseline firing rates, −9 ± 17% (± s.e.m.), $n = 13$, $P > 0.18$). Nearly all units (25 out of 28) showed a decrease in the suppression index (Fig. 4e), and in 10 out of 25 units the decrease was individually significant ($P < 0.05$ permutation test). The reduction of the suppression index was a result of SOM photo-hyperpolarization facilitating the response to large visual stimuli more than to small visual stimuli: the response ratio (the ratio of the firing rate in the illumination condition divided by firing rate in the control condition) increased with the size of the stimulus (Fig. 4f). Although the response to stimuli smaller or equal to the preferred size was not facilitated (−7 ± 7% (± s.e.m.), $P > 0.45$, paired signed rank test, $n = 28$; $\mathrm{FR}_{\mathrm{PS,\ CTRL}} = 4.9 \pm 1.5$ Hz, $\mathrm{FR}_{\mathrm{PS,\ LED}} = 4.6 \pm 1.9$ Hz, $P > 0.63$, paired signed rank test), the response to the stimuli larger than the preferred size was facilitated by 74 ± 19% ($P < 0.0011$, paired signed rank test, $n = 28$; Fig. 4f). This lack of facilitation of responses to smaller visual stimuli was not due to saturation (that is, a ceiling effect). In fact, firing rates to stimuli smaller or equal to the preferred one were consistently facilitated less than similar firing rates elicited by stimuli larger than the preferred one (Supplementary Fig. 8). The stronger impact of SOM photo-hyperpolarization on cortical responses to large stimuli is thus consistent with the preferential activation of SOMs by large stimuli (Fig. 1c). Hence, by inhibiting layer 2/3 neurons as a function of stimulus size, SOMs generate an inhibitory surround (Fig. 4g).

Discussion

This study describes a cortical circuit that significantly contributes to surround suppression of layer 2/3 cells, and identifies a specific type of

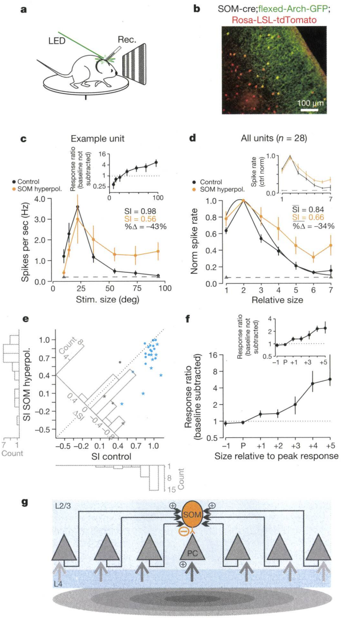


Figure 4 | SOMs contribute to size tuning of layer 2/3 pyramidal cells. **a**, Schematic of the experimental setup. **b**, Section of the visual cortex of a SOM-CRE;Rosa-LSL-tdTomato mouse injected with AAV-flexed-Arch-GFP. All SOM-CRE cells express tdTomato (red) and infected neurons also express Arch-GFP (green). **c**, Size tuning of an isolated unit in control conditions and during photo-hyperpolarization of SOMs. Dashed grey line, baseline firing rate. Inset, response ratio for this example unit. All error bars are s.e.m. **d**, Average peak-aligned and scaled size-tuning curves for 28 isolated single units (3 animals, 7 recording sessions). Black, control conditions; orange, SOM hyperpolarization. Inset, average peak-aligned and control-normalized size-tuning curve in the SOM-hyperpolarization condition. Note the lack of facilitation at the preferred or smaller size. All error bars are s.e.m. **e**, Scatter plot showing SI under control conditions (*x* axis) plotted against SI under SOM hyperpolarization (*y* axis) for each of the 28 units. Blue data points are units that are size tuned at baseline ($n = 24$, $\mathrm{SI} > 0$, $P < 0.05$). Grey data points are units that are not sized tuned at baseline. Stars are units that showed a significant reduction in SI ($n = 10$, $P < 0.05$). Histograms beside *x* and *y* axes show SI distribution under control and SOM hyperpolarization, respectively. The oblique histogram illustrates the distribution of changes in SI with SOM hyperpolarization. **f**, The ratio of the average response during SOM photo-hyperpolarization to the average response under control conditions plotted against stimulus size relative to peak. Same units as **d**. Inset, the same ratio without subtracting the baseline firing rate. **g**, Schematic illustration of the cortical circuit in layer 2/3 contributing to surround suppression. As a visual stimulus expands (larger stimuli are shown in lighter grey), recruitment of adjacent PCs increases SOM excitation through horizontal axons (horizontal arrows). Error bars on response ratios (**c** (inset), **f**) are estimates of s.e.m. (see Supplementary Methods for details).

inhibitory neuron, the SOM, as a key mediator of this phenomenon. This circuit is therefore likely to be involved in the contextual modulation of cortical responses to visual stimuli. The differential recruitment of pyramidal cells in superficial layers by ascending inputs and of SOMs by horizontal inputs underscores the fact that distinct neuron types are differentially integrated in the excitatory cortical circuit. These differences lead to different tuning properties, as highlighted here by the distinct size-tuning curves. Thus, although small stimuli efficiently drive layer 2/3 pyramidal cells through the activation of ascending vertical inputs, by summing activity in space via horizontal inputs, SOMs are preferentially driven by larger stimuli. As a consequence, the larger the stimulus, the stronger the SOM-mediated suppression of pyramidal cells.

An important questions is how SOMs increase their firing as a function of stimulus size if they suppress layer 2/3 pyramidal cells, that is, their main source of excitation. It is likely that the number of pyramidal cells recruited by the outer edge of the stimulus (an annulus that grows linearly with the diameter of the stimulus) more than compensate for the reduction in pyramidal-cell firing at the centre of the stimulus.

Photo-hyperpolarization of SOMs reduces but does not abolish surround suppression. Although this may partly be due to incomplete silencing of SOMs, surround suppression is probably also relayed to cortical layer 2/3 by earlier stages of visual processing[13–18], and other types of inhibitory neurons[22]or circuits[20] may also contribute to surround suppression.

The preferential recruitment of SOMs by horizontal excitatory projections is consistent with these projections having a role in size tuning[2]. In cortical layers with less extensive horizontal connectivity, size tuning may rely on different mechanisms[20,21] or may be entirely inherited from pre-cortical areas[13–18]. Importantly, because SOMs are tuned to the orientation of visual stimuli[36], they could account for the orientation dependence of surround suppression[8,9,37]. Furthermore, SOMs may respond differentially to specific stimulus properties, such as contrast, and thus also contribute to the contrast dependence of surround suppression[9,38,39]. It is likely that a connectivity pattern similar to what is described here may be present in other cortical areas as well, and thus contribute to suppressive surround in several sensory and non-sensory modalities.

METHODS SUMMARY

Experiments were performed in accordance with the regulations of the Institutional Animal Care and Use Committee of the University of California, San Diego. Mice were heterozygous for SOM-IRES-CRE (Jackson laboratory stock no. 013044) or PV-CRE (no. 008069) (except for the mice used for the experiments in Fig. 2a, b) and the reporter allele Rosa-LSL-tdTOMATO (Allen Institute line Ai9, Jackson Labs no. 007905). For Fig. 2a, b, mice were positive for Scnn1a-tg3-CRE (Jackson labs no. 009613) and crossed with the Gin (no. 003718) or B13 lines. For *in vivo* experiments, mice were implanted with a custom head plate and habituated to head fixation while running on a free-spinning circular treadmill. For targeted recording *in vivo*, tdTomato-expressing neurons were visualized by two-photon microscopy and contacted by a glass electrode containing Alexfluor 488. Extracellular unit recording was carried out using 16-channel silicon probes (Neuronexus). Single units were isolated using custom spike sorting software (Kleinfeld laboratory). We conditionally expressed ChR2 by *in utero* electroporation (for layer 2/3) or through a CRE-dependent adeno-associated virus (AAV) in Scnn1a-tg3-CRE (for layer 4). Arch or eNpHR were expressed via CRE-dependent AAVs in SOM- and PV-IRES-CRE mice. Visual stimuli were generated by custom software (Psych Toolbox) and presented on a gamma-corrected liquid crystal display monitor 15 cm from the mouse. Photostimulation *in vivo* was carried out using fibre-coupled light-emitting diodes (LEDs; Doric lenses). Photostimulation *in vitro* was carried out using a combination of fibre-coupled LEDs, or LEDs mounted and coupled to an epifluorescence microscope (Olympus BX51). eNpHR was activated by a shuttered arc lamp. Slice preparation and intracellular recording followed previous protocols. Data acquisition, visual stimulation and statistical analysis was carried out using Igor Pro and Matlab.

Received 1 March; accepted 17 August 2012.

1. Allman, J., Miezin, F. & McGuinness, E. Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* **14,** 105–126 (1985).
2. Angelucci, A. & Bressloff, P. C. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog. Brain Res.* **154,** 93–120 (2006).
3. Gilbert, C. D., Das, A., Ito, M., Kapadia, M. & Westheimer, G. Spatial integration and cortical dynamics. *Proc. Natl Acad. Sci. USA* **93,** 615–622 (1996).
4. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.* **28,** 229–289 (1965).
5. Blakemore, C. & Tobin, E. A. Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp. Brain Res.* **15,** 439–440 (1972).
6. Nelson, J. I. & Frost, B. J. Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res.* **139,** 359–365 (1978).
7. DeAngelis, G. C., Freeman, R. D. & Ohzawa, I. Length and width tuning of neurons in the cat's primary visual cortex. *J. Neurophysiol.* **71,** 347–374 (1994).
8. Knierim, J. J. & van Essen, D. C. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* **67,** 961–980 (1992).
9. Levitt, J. B. & Lund, J. S. Contrast dependence of contextual effects in primate visual cortex. *Nature* **387,** 73–76 (1997).
10. Lamme, V. A. The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15,** 1605–1615 (1995).
11. Dobbins, A., Zucker, S. W. & Cynader, M. S. Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature* **329,** 438–441 (1987).
12. Mareschal, I. & Shapley, R. M. Effects of contrast and size on orientation discrimination. *Vision Res.* **44,** 57–67 (2004).
13. Solomon, S. G., Lee, B. B. & Sun, H. Suppressive surrounds and contrast gain in magnocellular-pathway retinal ganglion cells of macaque. *J. Neurosci.* **26,** 8715–8726 (2006).
14. Alitto, H. J. & Usrey, W. M. Origin and dynamics of extraclassical suppression in the lateral geniculate nucleus of the macaque monkey. *Neuron* **57,** 135–146 (2008).
15. Zhang, Y., Kim, I.-J., Sanes, J. R. & Meister, M. The most numerous ganglion cell type of the mouse retina is a selective feature detector. *Proc. Natl Acad. Sci. USA* advance online publication, doi: 10.1073/pnas.1211547109 (13 August 2012).
16. Murphy, P. C. & Sillito, A. M. Corticofugal feedback influences the generation of length tuning in the visual pathway. *Nature* **329,** 727–729 (1987).
17. Sceniak, M. P., Chatterjee, S. & Callaway, E. M. Visual spatial summation in macaque geniculocortical afferents. *J. Neurophysiol.* **96,** 3474–3484 (2006).
18. Bonin, V., Mante, V. & Carandini, M. The suppressive field of neurons in lateral geniculate nucleus. *J. Neurosci.* **25,** 10844–10856 (2005).
19. Ozeki, H. *et al.* Relationship between excitation and inhibition underlying size tuning and contextual response modulation in the cat primary visual cortex. *J. Neurosci.* **24,** 1428–1438 (2004).
20. Bolz, J. & Gilbert, C. D. Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature* **320,** 362–365 (1986).
21. Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D. & Ferster, D. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62,** 578–592 (2009).
22. Haider, B. *et al.* Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. *Neuron* **65,** 107–121 (2010).
23. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28,** 7520–7536 (2008).
24. Van den Bergh, G., Zhang, B., Arckens, L. & Chino, Y. M. Receptive-field properties of V1 and V2 neurons in mice and macaque monkeys. *J. Comp. Neurol.* **518,** 2051–2070 (2010).
25. Margrie, T. W. *et al.* Targeted whole-cell recordings in the mammalian brain *in vivo*. *Neuron* **39,** 911–918 (2003).
26. Kawaguchi, Y. & Kubota, Y. GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* **7,** 476–486 (1997).
27. McCormick, D. A., Connors, B. W., Lighthall, J. W. & Prince, D. A. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* **54,** 782–806 (1985).
28. Taniguchi, H. *et al.* A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71,** 995–1013 (2011); erratum **72,** 782–806 (2011).
29. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8,** 1263–1268 (2005).
30. Petreanu, L., Mao, T., Sternson, S. M. & Svoboda, K. The subcellular organization of neocortical excitatory connections. *Nature* **457,** 1142–1145 (2009).
31. Adesnik, H. & Scanziani, M. Lateral competition for cortical space by layer-specific horizontal circuits. *Nature* **464,** 1155–1160 (2010).
32. Wang, Q. & Burkhalter, A. Area map of mouse visual cortex. *J. Comp. Neurol.* **502,** 339–357 (2007).
33. Kapfer, C., Glickfeld, L. L., Atallah, B. V. & Scanziani, M. Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nature Neurosci.* **10,** 743–753 (2007).
34. Chow, B. Y. *et al.* High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* **463,** 98–102 (2010).
35. Atasoy, D., Aponte, Y., Su, H. H. & Sternson, S. M. A. FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J. Neurosci.* **28,** 7025–7030 (2008).
36. Ma, W. P. *et al.* Visual representations by cortical somatostatin inhibitory neurons–selective but with weak and delayed responses. *J. Neurosci.* **30,** 14371–14379 (2010).
37. Cavanaugh, J. R., Bair, W. & Movshon, J. A. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* **88,** 2547–2556 (2002).
38. Kapadia, M. K., Westheimer, G. & Gilbert, C. D. Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc. Natl Acad. Sci. USA* **96,** 12073–12078 (1999).
39. Sceniak, M. P., Ringach, D. L., Hawken, M. J. & Shapley, R. Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neurosci.* **2,** 733–739 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to J. Evora for the reconstruction of SOMs and technical assistance. We thank C. Niell and M. Stryker for providing expertise and sharing code used at the initial stages of this project; S. Olsen for providing the firing rates of part of the units isolated under anaesthesia; P. Abelkop and A. Linder for technical assistance; and J. Isaacson and members of the Scanziani laboratory for helpful advice. H.A. was supported by the Helen Hay Whitney Foundation and the Howard Hughes Medical Institute (HHMI). W.B. and M.S. were supported by the HHMI, the Gatsby charitable foundation and US National Institute of Health grant NS069010.

Author Contributions H.A. and M.S. designed the study. H.A. conducted all experiments. W.B. conducted all *in vivo* data analysis and spike sorting. H.T. and Z.J.H. generated the SOM-IRES-CRE mice. M.S. and H.A. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.S. (massimo@ucsd.edu).

LETTER

doi:10.1038/nature11511

Unexpectedly large mass loss during the thermal pulse cycle of the red giant star R Sculptoris

M. Maercker[1,2], S. Mohamed[3], W. H. T. Vlemmings[4], S. Ramstedt[2], M. A. T. Groenewegen[5], E. Humphreys[1], F. Kerschbaum[6], M. Lindqvist[4], H. Olofsson[4], C. Paladini[6], M. Wittkowski[1], I. de Gregorio-Monsalvo[7] & L.-A. Nyman[7]

The asymptotic-giant-branch star R Sculptoris is surrounded by a detached shell of dust and gas[1,2]. The shell originates from a thermal pulse during which the star underwent a brief period of increased mass loss[3,4]. It has hitherto been impossible to constrain observationally the timescales and mass-loss properties during and after a thermal pulse—parameters that determine the lifetime of the asymptotic giant branch and the amount of elements returned by the star. Here we report observations of CO emission from the circumstellar envelope and shell around R Sculptoris with an angular resolution of 1.3″. What was previously thought to be only a thin, spherical shell with a clumpy structure is revealed to also contain a spiral structure. Spiral structures associated with circumstellar envelopes have been previously seen, leading to the conclusion that the systems must be binaries[5–8]. Combining the observational data with hydrodynamic simulations, we conclude that R Sculptoris is a binary system that underwent a thermal pulse about 1,800 years ago, lasting approximately 200 years. About 3×10^{-3} solar masses of material were ejected at a velocity of $14.3\ \mathrm{km\ s^{-1}}$ and at a rate around 30 times higher than the pre-pulse mass-loss rate. This shows that about three times more mass was returned to the interstellar medium during and immediately after the pulse than previously thought.

The detached shell around R Sculptoris was observed in CO($J = 3 - 2$) emission at 345 GHz using the Atacama Large Millimeter/submillimeter Array (ALMA) during Cycle 0 operations (Fig. 1 and Supplementary Information). The data clearly show the well-centred detached shell with a radius of 18.5″, and reveal a spiral structure extending from the central star outwards to the shell. Previous observations of R Sculptoris show structure in the form of clumps. However, this was interpreted as clumpy material within the shell itself, and not as a structure interior to the shell[2].

Until now no clear signs of binary companions have been observed in any detached shell sources (with a possible exception for the detached shell around TT Cygnus[9]). The observed structure around R Sculptoris, however, indicates the presence of a companion, shaping the stellar wind into a spiral shell structure[8]. Smoothed particle hydrodynamics (SPH) models show that a wide binary companion can have a significant effect in the shaping of the wind, leading to elliptical and spiral structures (for example, as observed in the case of the envelope of AFGL 3068)[5,6].

The observed shapes of the circumstellar envelopes around binary asymptotic-giant-branch stars depend on the physical parameters of the binary system (such as separation and mass ratio[10]), the density contrasts imprinted on the wind, the temperatures in the circumstellar envelope, the viewing angle, and the chemistry and excitation of the gas[11]. The temporal variations of the mass-loss rate and the expansion velocity further affect the structure of the circumstellar envelope. Hence, the observed spiral structure and detached shell allow us to measure these important properties, and to link them directly to the thermal pulse.

Any change in the expansion velocity of the stellar wind will affect the spacing between the spiral windings. In Fig. 2 the spiral can be followed from the central star out to the detached shell over about five windings. The 2.5 inner windings have a nearly constant spacing, with an average distance of 2.6″, implying an essentially constant expansion velocity during the last 2.5 orbital periods. The expansion velocity of the present-day wind[3] of R Sculptoris gives an orbital period of $t_{\mathrm{orb}} = 350$ years. The position angle and radius of the observed emission then allow us to derive the velocity evolution of the stellar wind

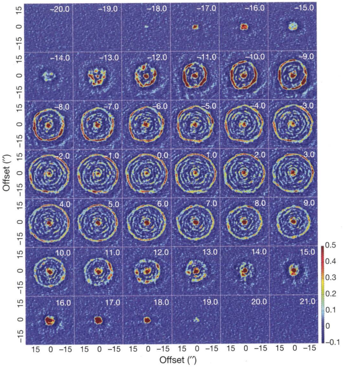


Figure 1 | ALMA Early Science observations of the CO($J = 3 - 2$) emission from the asymptotic-giant-branch star R Sculptoris. The figure shows the emission in the different velocity channels. The colour scale gives the flux in Jy per beam. The stellar velocity with respect to the local standard of rest (LSR) is $v_{\mathrm{LSR}} = -19\ \mathrm{km\ s^{-1}}$. The numbers in the top right corners indicate the velocity in kilometres per second with respect to the stellar velocity. The spherical detached shell appears as a ring in the individual velocity channels, with its largest extent at the stellar velocity. The shell is clearly visible at 18.5″ at the stellar v_{LSR}, as well as a spiral structure connecting the central star with the detached shell. The structure can be traced through all velocity channels.

---

[1]European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching, Germany. [2]Argelander Institute for Astronomy, University of Bonn, Auf dem Hügel 71, 53121 Bonn, Germany. [3]South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, Western Cape, South Africa. [4]Onsala Space Observatory, Department of Earth and Space Sciences, Chalmers University of Technology, SE-43992 Onsala, Sweden. [5]Royal Observatory of Belgium, Ringlaan 3, 1180 Brussels, Belgium. [6]University of Vienna, Department of Astrophysics, Türkenschanzstrasse 17, 1180 Wien, Austria. [7]Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile.

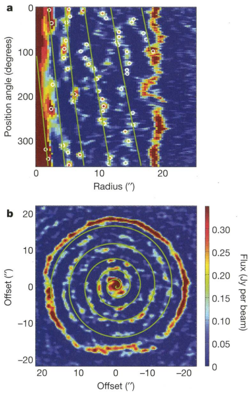


Figure 2 | The CO($J = 3 - 2$) emission at the stellar v_{LSR} of R Sculptoris. **a**, Position angle versus radius based on the stellar v_{LSR} image. The position angle starts at due North and increases counter-clockwise. **b**, The stellar v_{LSR} image. The colour scale gives the flux in Jy per beam. The green lines show linear fits to the emission peaks (white circles) in **a**. The first 2.5 windings are nearly parallel, with a constant separation of 2.6″ ± 0.07″, indicating that the expansion velocity has been constant (on average) for the last 2.5 binary periods. The present-day expansion velocity is estimated to be 10.5 km s^{-1}, giving a binary period of 350 years. The linear fits can hence be translated directly into a velocity evolution (Fig. 3). The corresponding spiral is plotted on top of the stellar v_{LSR} image (**b**). Deviations from a perfect spiral are of the order of ±1.5 km s^{-1}, indicating small velocity variations over periods of 50 years. Partial spiral arms and arcs connecting the third and fourth windings show a larger variation in the wind velocity during these orbital periods on a timescale of about 100 years.

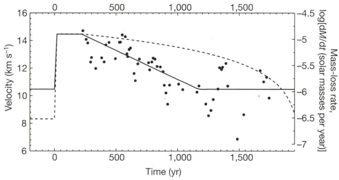


Figure 3 | The velocity and mass-loss rate evolution of the stellar wind around R Sculptoris. The solid and dashed lines show the velocity and mass-loss rate as a function of time, respectively. The points correspond to the expansion velocities of the emission peaks in Fig. 2a, assuming a binary period of 350 years. The figure shows the evolution of the velocity and mass-loss rate since the onset of the last thermal pulse. The velocity profile is a fit to the data points, whereas the mass-loss rate profile is constrained by the pre-pulse, pulse and present-day mass-loss rates. The profiles are used as input to the SPH models. The shape of the mass-loss rate profile is chosen to be consistent with the observations. The overall velocity fits the predictions from theoretical models of thermal pulses well. However, velocity variations of ±1.5 km s^{-1} are apparent throughout the evolution since the last pulse, whereas theoretical models only predict significant variations in the expansion velocity less than 200 years after the pulse[4]. Theoretical models predict an increasing widening of the spiral proportional to the local sound speed, which may explain at least part of the velocity variation[8].

from R Sculptoris from the star out to the detached shell (Fig. 3, and Supplementary Information). The derived evolution of the expansion velocity since the last thermal pulse is consistent with models of thermal pulses[4]. However, the observed emission implies variations in the expansion velocity of ±1.5 km s^{-1} on timescales of a few hundred years. Observed partial spiral windings and arcs, as well as emission at velocities up to 19 km s^{-1}, indicate brief periods of even larger velocity variations.

A spherically symmetric, detached shell can still be created in a binary system where the asymptotic-giant-branch star is undergoing a thermal pulse, owing to the brief increase in mass-loss rate and expansion velocity. Collision with the surrounding, slower material will then shape the wind into a symmetric shell structure. The post-pulse mass loss leaves behind a spiral structure that connects the detached shell with the central star. Assuming a spherically symmetric expanding detached shell gives a shell expansion velocity of $v_{\mathrm{sh}} = 14.3$ km s^{-1} and a shell radius of $R_{\mathrm{sh}} = 18.5''$ (see Supplementary Information).

The present expansion velocity and size of the shell around R Sculptoris put the upper limit to the end of the thermal pulse at 1,800 years ago. With a binary period of 350 years we would expect to see around five windings since the pulse, consistent with the observed spiral. A decelerated shell would imply a shorter time since the thermal pulse, and hence a shorter binary period or fewer spiral windings. The sweeping-up of material from the pre-pulse wind could cause such a deceleration of the shell. However, although a slight decrease in the expansion velocity of the detached shell is possible, we find no evidence of a noticeable decrease. This is contrary to the current theory of how detached shells are formed during thermal pulses[3,4]. Also, no material is likely to have piled onto the shell because of the post-pulse mass loss.

Recent images of thermal dust emission show the detached shell, as well as a more distant, and spatially distinct, region of interaction with the interstellar material[12], indicating the presence of a stellar wind from R Sculptoris before the pulse. Collision with a previous, slower wind is required to prevent the thin shell from quickly diffusing[4,13]. The average CO line intensity in the area surrounding the shell sets an upper limit to the pre-pulse mass-loss rate, and its ratio to the average CO line intensity in the shell suggests an increase in mass-loss rate during the thermal pulse by a factor of more than ten. The total shell mass is essentially only due to the mass lost during the formation of the shell. An estimated total gas mass in the shell of 2.5×10^{-3} solar masses (ref. 3) gives a thermal pulse mass-loss rate of between 7×10^{-6} solar masses per year and 2.5×10^{-5} solar masses per year (see Supplementary Information), and implies a pre-pulse mass-loss-rate of less than 10^{-6} solar masses per year. The present-day mass-loss rate is estimated to be 3×10^{-7} solar masses per year (ref. 3), that is, a factor of about 30 lower than during the pulse. This general evolution of the mass-loss rate is consistent with stellar evolution models; however, the ratio between the derived pulse and pre-pulse mass-loss-rate is significantly higher than found in the models[4].

To further constrain the mass-loss-rate evolution of R Sculptoris, we modelled the system with a modified version of the GADGET-2 SPH code[14], including detailed radiative cooling[15]. The modelled system successfully forms a detached shell, including the observed spiral structure (see Supplementary Video). The modelled density, temperature and velocity structures of the SPH model are then used as input in the three-dimensional radiative transfer code LIME[16]. The global morphology of the modelled system closely resembles that of the observations, and the brightness distribution reproduces the observed intensities well (Fig. 4). We effectively constrain the mass-loss-rate evolution

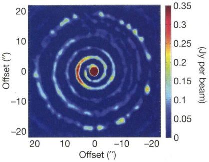


Figure 4 | LIME radiative transfer model of the circumstellar structure around R Sculptoris. The model is based on the results from the SPH models at the stellar velocity. The results of the radiative transfer model have been processed by the 'simdata' task in CASA (Common Astronomy Software Applications) using the ALMA Cycle 0 compact configuration specifications. The colour scale gives the flux in Jy per beam. The overall model intensities match the observed intensities well, while variations in the intensity contrast between the spiral windings and interwinding material indicate a more complicated mass-loss rate variation. The inclination angle of the binary system to the line of sight is 90°. See Supplementary Information for an animation of the SPH models of R Sculptoris.

throughout the thermal pulse to the present time (Fig. 3). Assuming an inter-pulse time of 50,000 years (typical for stars of about 1−4 solar masses; ref. 17) and our derived mass-loss rate evolution, 10% of the mass lost between two subsequent pulses is expelled during the thermal pulse, and 40% is lost during the first 1,800 years after the pulse. The pulse and immediate post-pulse phases are therefore dominant in the formation and chemical enrichment of the circumstellar envelope.

The chemical content of the expelled material depends critically on the physical properties of the pulses (for example, pulse duration and interpulse mass-loss rate). The duration of the pulse limits the time for nucleosynthesis to occur inside the star[18], whereas the mass-loss rate between pulses limits the number of thermal pulses an asymptotic-giant-branch star will experience[17]. These properties will affect the stellar yields of new elements returned to the interstellar medium, as well as eventually lead to the termination of the asymptotic-giant-branch phase. The observations presented here directly constrain these important physical parameters throughout the thermal-pulse cycle. In essence, it is the observed spiral structure that allows us to verify model results observationally, and refine our knowledge of thermal pulses and late stellar evolution.

Received 20 April; accepted 10 August 2012.

1. Olofsson, H., Eriksson, K. & Gustafsson, B. SEST CO (J = 1 - 0) observations of carbon-rich circumstellar envelopes. *Astron. Astrophys.* **196,** L1–L4 (1988).
2. Olofsson, H., Maercker, M., Eriksson, K., Gustafsson, B. & Schöier, F. High-resolution HST/ACS images of detached shells around carbon stars. *Astron. Astrophys.* **515,** A27 (2010).
3. Schöier, F. L., Lindqvist, M. & Olofsson, H. Properties of detached shells around carbon stars. Evidence of interacting winds. *Astron. Astrophys.* **436,** 633–646 (2005).
4. Mattsson, L., Höfner, S. & Herwig, F. Mass loss evolution and the formation of detached shells around TP-AGB stars. *Astron. Astrophys.* **470,** 339–352 (2007).
5. Mastrodemos, N. & Morris, M. Bipolar pre-planetary nebulae: hydrodynamics of dusty winds in binary systems. II. Morphology of the circumstellar envelopes. *Astrophys. J.* **523,** 357–380 (1999).
6. Mauron, N. & Huggins, P. J. Imaging the circumstellar envelopes of AGB stars. *Astron. Astrophys.* **452,** 257–268 (2006).
7. Dinh-V-Trung & Lim, J. Tracing the asymmetry in the envelope around the carbon star CIT 6. *Astrophys. J.* **701,** 292–297 (2009).
8. Kim, H. & Taam, R. E. Probing substellar companions of asymptotic giant branch stars through spirals and arcs. *Astrophys. J.* **744,** 136 (2012).
9. Olofsson, H. *et al.* A high-resolution study of episodic mass loss from the carbon star TT Cygni. *Astron. Astrophys.* **353,** 583–597 (2000).
10. De Marco, O. The origin and shaping of planetary nebulae: putting the binary hypothesis to the test. *Publ. Astron. Soc. Pacif.* **121,** 316–342 (2009).
11. Politano, M. & Taam, R. E. The incidence of non-spherical circumstellar envelopes in asymptotic giant branch stars. *Astrophys. J.* **741,** 5 (2011).
12. Cox, N. L. J. *et al.* A far-infrared survey of bowshocks and detached shells around AGB stars and red supergiants. *Astron. Astrophys.* **537,** A35 (2012).
13. Steffen, M. & Schönberner, D. On the origin of thin detached gas shells around AGB stars. Insights from time-dependent hydrodynamical simulations. *Astron. Astrophys.* **357,** 180–196 (2000).
14. Springel, V. The cosmological simulation code GADGET-2. *Mon. Not. R. Astron. Soc.* **364,** 1105–1134 (2005).
15. Mohamed, S. & Podsiadlowski, P. in *15th European Workshop on White Dwarfs* (eds R. Napiwotzki & M. R. Burleigh) 397–400 (Astronomical Society of the Pacific Conference Series 372, 2007).
16. Brinch, C. & Hogerheijde, M. R. LIME—a flexible, non-LTE line excitation and radiation transfer method for millimeter and far-infrared wavelengths. *Astron. Astrophys.* **523,** A25 (2010).
17. Karakas, A. & Lattanzio, J. C. Stellar models and yields of asymptotic giant branch stars. *Publ. Astron. Soc. Aust.* **24,** 103–117 (2007).
18. Busso, M., Gallino, R. & Wasserburg, G. J. Nucleosynthesis in asymptotic giant branch stars: relevance for galactic enrichment and solar system formation. *Annu. Rev. Astron. Astrophys.* **37,** 239–309 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements This paper makes use of ALMA data from project no. ADS/JAO.ALMA#2011.0.00131.S. ALMA is a partnership of ESO (representing its member states), the NSF (USA) and NINS (Japan), together with the NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. We gratefully acknowledge the technical expertise and assistance provided by the Spanish Supercomputing Network (Red Espanola de Supercomputacion), as well as the use of the LaPalma Supercomputer, located at the Instituto de Astrofisica de Canarias. F.K. acknowledges funding by the Austrian Science Fund FWF under project numbers P23586-N16 and I163-N16. C.P. acknowledges funding by the Austrian Science Fund FWF under project number P23006-N16.

Author Contributions M.M. planned the project, prepared and submitted the proposal, analysed the data, and wrote the manuscript. S.M. was involved in project preparation, data interpretation, did the SPH modelling, and commented on the manuscript. W.V. was involved in project planning, proposal preparation, data reduction and analysis, radiative transfer modelling, and commented on the manuscript. S.R. was involved in project planning, data analysis, and commented on the manuscript. The remaining authors were involved in the project preparation, science discussion, and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.M. (mmaercke@eso.org).

doi:10.1038/nature11562

Probing graphene grain boundaries with optical microscopy

Dinh Loc Duong[1,2]*, Gang Hee Han[2]*, Seung Mi Lee[3], Fethullah Gunes[2], Eun Sung Kim[2], Sung Tae Kim[2], Heetae Kim[2], Quang Huy Ta[2], Kang Pyo So[2], Seok Jun Yoon[2], Seung Jin Chae[1,2], Young Woo Jo[2], Min Ho Park[4], Sang Hoon Chae[2], Seong Chu Lim[2], Jae Young Choi[2,5] & Young Hee Lee[1,2]

Grain boundaries in graphene are formed by the joining of islands during the initial growth stage, and these boundaries govern transport properties and related device performance[1,2]. Although information on the atomic rearrangement at graphene grain boundaries can be obtained using transmission electron microscopy[3,4] and scanning tunnelling microscopy[2,5–8], large-scale information regarding the distribution of graphene grain boundaries is not easily accessible. Here we use optical microscopy to observe the grain boundaries of large-area graphene (grown on copper foil) directly, without transfer of the graphene. This imaging technique was realized by selectively oxidizing the underlying copper foil through graphene grain boundaries functionalized with O and OH radicals generated by ultraviolet irradiation under moisture-rich ambient conditions: selective diffusion of oxygen radicals through OH-functionalized defect sites was demonstrated by density functional calculations. The sheet resistance of large-area graphene decreased as the graphene grain sizes increased, but no strong correlation with the grain size of the copper was revealed, in contrast to a previous report[9]. Furthermore, the influence of graphene grain boundaries on crack propagation (initialized by bending) and termination was clearly visualized using our technique. Our approach can be used as a simple protocol for evaluating the grain boundaries of other two-dimensional layered structures, such as boron nitride and exfoliated clays.

Optical microscopy is an important tool for characterizing graphene at large scales. Although graphene is only one atom thick, a given area can have more than one layer growing on it; areas with differing numbers of layers can be distinguished by their contrast difference[10]. Optical birefringence from transferred graphene covered by a liquid crystal has been used to visualize graphene grain boundaries (GGBs), but the grain boundaries of the copper substrate were visualized, and looked like GGBs[9]. This last observation contradicts previous reports based on scanning electron microscopy (SEM) in which GGBs and copper grain boundaries were found to be independent[11]. Direct characterization of GGBs on copper foil is required. Our work examines graphene grown directly on copper without the need for a transfer process, which often leads to undesired artefacts, such as additional traces from copper substrates, wrinkles and/or cracks[12].

The primary objective of this Letter is to visualize grain boundaries in large-area graphene, and distinguish them from grain boundaries of the copper substrate, using an optical microscope. This approach is based on the robust oxidization of copper foil at room temperature via the selective diffusion of O and OH radicals through the GGBs (Fig. 1a). These radicals are generated using ultraviolet irradiation

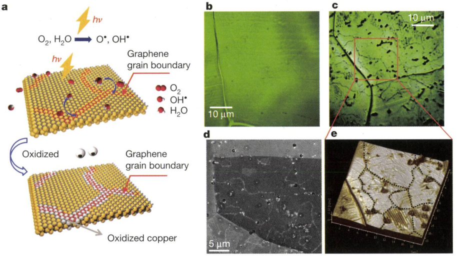


Figure 1 | Observation of graphene grain boundaries (GGBs) after ultraviolet exposure under moisture-rich ambient conditions. a, Diagram of the ultraviolet treatment of a graphene/Cu sample. The copper under the GGBs was oxidized by radicals; the lines of oxidized copper were broadened during continuing oxidation and thereby became visible using an optical microscope. **b**, **c**, Optical images of graphene/Cu before (**b**) and after (**c**) oxidation. The GGBs were visible after oxidation. **d**, SEM image of oxidized graphene/Cu. The GGBs (coincident with oxidized copper seen as white dotted lines) intersected the Cu grain boundary (between the dark central region and the surrounding grey region). **e**, AFM force image of the position marked by a red square in **c**.

---

[1] Sungkyunkwan Advanced Institute of Nanotechnology, Sungkyunkwan University, Suwon 440-746, South Korea. [2] Department of Energy Science, BK21 Physics Division, Graphene Center, Sungkyunkwan University, Suwon 440-746, South Korea. [3] Center for Nanocharacterization, Korea Research Institute of Standards and Science, Daejeon 305-340, South Korea. [4] School of Advanced Materials Science and Engineering, Sungkyunkwan University, Suwon 440-746, South Korea. [5] Graphene Center at Samsung Advanced Institute of Technology (SAIT), PO Box 111, Suwon 440-600, South Korea.
*These authors contributed equally to this work.

under moisture-rich ambient conditions. The GGBs are very narrow, but the width of the oxidized copper foil beneath the defective GGBs is increased by the continuous supply of oxidizing radicals through the grain boundary. In addition, oxidized copper occupies a larger volume than copper, thereby allowing the visualization of the GGBs using an optical microscope (Fig. 1a).

In general, GGBs are not visible using optical microscopy because of the nanoscale width of the boundaries[2–8]. Although copper grain boundaries were clearly observed in the optical microscopy images because of their large width, GGBs were not visible (Fig. 1b). The CVD-grown large-area graphene on the Cu foil was used directly without transfer in our study (see Methods). After the graphene/Cu sample was irradiated by ultraviolet light under moisture-rich ambient conditions, the GGBs were visible as thin dark lines. Some of the GGB lines crossed over the Cu grain boundaries, demonstrating the formation of GGBs regardless of the location of the Cu grain boundaries, which is consistent with previous investigations using SEM[11]. Similar GGBs were also observed as white lines in scanning electron microscopy (SEM) images (Fig. 1d). The copper grain boundaries were clearly distinguishable from the GGBs on the basis of their channelling contrast levels (that is, different orientations of the copper crystals result in different brightnesses; Supplementary Fig. 1). It is interesting to note the different thicknesses of the GGB lines, which implies that the nature of the defects in GGBs is complicated. Higher oxygen content was observed at the GGBs than within the graphene grains (Supplementary Fig. 2). The GGBs in the optical image also matched well with those present in the non-contact force image obtained using tapping-mode atomic force microscopy (AFM; Fig. 1e).

The GGBs visible in the optical image (Fig. 2a) were also confirmed using confocal Raman mapping. The CVD-grown graphene/Cu exhibited a large intensity difference between the G′ and G bands (a high G′/G ratio) with no appreciable D-band (Fig. 2b), thereby demonstrating a high-quality monolayer of graphene. With ultraviolet treatment, the D-band developed both at the grain boundary and within the grain, but occurred more strongly at the grain boundary. This developed D-band intensity was the result of the increased oxygen content, which increased from 2.8 to 14.2 at.%, as confirmed using X-ray photoelectron spectroscopy analysis (Supplementary Fig. 3). The blueshift of the G-band (1,590 to 1,597 cm^{-1}) and the G′-band (2,700 to 2,717 cm^{-1}) in addition to the D-band development inside the grain after ultraviolet treatment is indicative of graphene oxidation[13,14]. Interestingly, at the grain boundary, both the G-band and the G′-band peaks were redshifted compared with those within the grain, respectively from 1,597 to 1,594 cm^{-1} and from 2,717 to 2,688 cm^{-1}. This redshift may be explained as follows: in addition to the oxidation of the pristine graphene at the grain boundary, the oxidized copper formed under the GGB has a greater volume than the pristine copper, and this increased volume forced the uppermost graphene to experience a tensile strain[15,16]. The mappings in D-band, G-band and G′-band (Fig. 2c–e) matched well with the GGBs observed using optical microscopy. We emphasize here that GGBs cannot be observed using confocal Raman mapping of graphene unless prior ultraviolet treatment has been applied (Supplementary Fig. 4).

The observation of GGBs using optical microscopy after ultraviolet treatment is possible owing to the morphological changes in the sample. Three types of line profile were detected using topological AFM before and after ultraviolet irradiation (Fig. 3a): (1) the GGBs; (2) a wrinkle caused by a GGB; and (3) a Cu grain boundary. The step-like GGBs (with a step height of 2.5 nm) were visualized more clearly after ultraviolet treatment (Fig. 3c top), which enhanced the step height to approximately 15 nm because of the expanded volume of oxidized Cu at the boundary. The width of the oxidized boundary increased to approximately 500–600 nm, which is sufficiently large to be observed using an optical microscope. The second type of GGB line overlapped the wrinkle line and was approximately 5 nm high[12,17]. This observation is consistent with a previous theoretical report that predicted wrinkles might be formed from GGBs[18]. After ultraviolet irradiation, the GGB was distinguishable because of its increased height (Fig. 3c middle). This effect is similar to the previous example but the wrinkle line remained unchanged. On the other hand, the Cu grain boundary was initially grooved to a depth of 15 nm, and this depth was not altered by oxidation (Fig. 3c bottom). This observation supports previous reports of growing graphene sliding over the Cu grain boundary without creating defects[11] but contrasts with a report of strongly correlated GGBs and Cu grain boundaries[9]. Our results provide further evidence of the clear distinction between GGBs and Cu grain boundaries. Further analysis of the

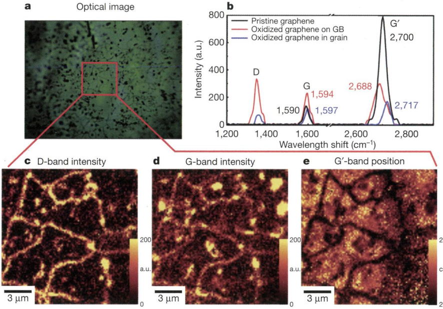


Figure 2 | Confocal Raman mapping of GGBs. **a**, **b**, Optical image (**a**) and Raman spectra (**b**) of graphene/Cu. The blueshift of the G- and G′-bands in addition to the development of the D-band within the grain provide evidence of graphene oxidation. The redshift of the G- and G′-band peaks on the grain boundary compared with those within the grain was caused by the tensile strain developed in the graphene owing to the expanded underlying oxidized copper. **c**–**e**, Maps of the intensities of the D-band (**c**) and G-band (**d**), and the shift of the G′-band (**e**).

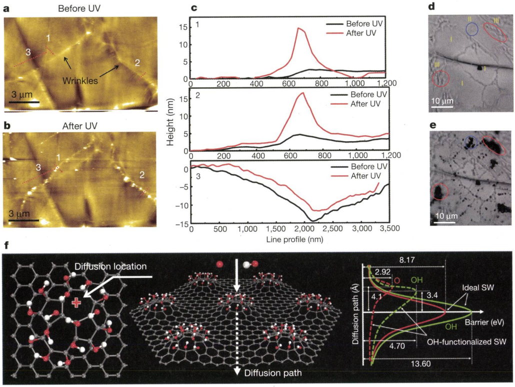


Figure 3 | Height profiles of various topological GGBs, and the oxidation mechanism. a, b, Topological images of various GGBs; (1) GGB, (2) a wrinkle caused by a GGB, and (3) a Cu grain boundary. **c,** Height profiles of the selected positions 1–3. **d, e,** Optical images of graphene/Cu during early growth (for 5 s) before (**d**) and after (**e**) oxidation. Three distinct regions are identified: (I) a typical graphene island, (II) a highly defective graphene region, and (III) Cu foil with incomplete graphene growth. **f,** Density functional calculations for diffusion of radicals through a heptagon of OH-functionalized Stone–Wales defects (SW). Stone–Wales defects were enlarged and deactivated by the adsorbed radicals.

morphology of oxidized copper and graphene after ultraviolet irradiation is shown in Supplementary Figs 5 and 6.

A GGB is typically formed by the joining of adjacent graphene islands that were initiated by nucleation seeds. To determine the relevance of our approach to observing the initial growth stage of graphene islands, graphene was grown for only 5 s using CVD conditions similar to those described in Methods. Three distinct regions near a Cu grain boundary line were observed under these conditions (Fig. 3d): a typical graphene island (feature I), a highly defective graphene region (feature II), and Cu foil with incomplete graphene growth (feature III). The black spot in the middle of the grain (Fig. 3d) is a nucleation seed that has also previously been observed[12]. The graphene/Cu shown in Fig. 3d was then subjected to ultraviolet irradiation under humid conditions; in Fig. 3e we show the area corresponding to Fig. 3d.It is interesting to note the occurrence of divergent grain boundary lines within the island centred at the nucleation seed, which is in good agreement with previous confocal Raman mapping and TEM observations[6,12]. The highly defective graphene region and the graphene incompletely covering the Cu foil were readily oxidized and converted into black spots by the facilitation of radical diffusion. Thus, microstructures developed during the initial growth stage can be clearly observed using a simple optical microscope.

The selective oxidation of Cu foil through the GGBs is dependent on the oxidation conditions. For example, the humidity level during ultraviolet irradiation is crucial in developing clear oxidation patterns. At low humidity levels (less than 25%), the grain boundaries were only partially visible (Supplementary Fig. 7a–d). Even when the oxidation time was prolonged to 30 min at these low humidity levels, GGBs were not observed. Pure ozone could not oxidize the graphene or the copper substrate at room temperature, as reported previously[19,20]. The presence of H_2O was necessary to create OH radicals under ultraviolet irradiation[21,22]; these radicals are necessary to functionalize graphene defects to further facilitate copper oxidation (see simulations in Methods). At a humidity level of 25%, the black spots began to appear. These black spots, which may be impurities introduced during the preparation of the Cu foil, acted as nucleation seeds[12]. The black spots could also be attributable to defects in the graphene. At higher humidity levels, the size of the black spots became larger, and the GGB lines were clearly visible. However, the size of the black spots and the GGB lines did not increase at a greater humidity level (more than 66%) or with longer oxidation times (Supplementary Fig. 5e–h). It is important to note that in addition to the oxidation of the graphene (by forming epoxide, hydroxyl and carboxyl groups particularly at the defect sites), the copper foil was also oxidized to give exclusively $Cu(OH)_2$ (Supplementary Fig. 8).

We performed density functional calculations to understand the oxidation mechanism associated with GGB defects (Fig. 3f, Supplementary Figs 9 and 10, Supplementary Table 1). The Stone–Wales defect experimentally observed at GGBs[3,4] was modelled. The height of the barrier experienced by species diffusing through the graphene was consistently reduced at the functionalized Stone–Wales defects, in particular, those functionalized by OH. It was interesting to see that the O radical had a smaller diffusion barrier height than the OH radical. In other words, the role of OH radicals was to reduce the diffusion barrier height by functionalizing defects. Also, the absolute values of the diffusion barrier height could be lowered by taking into account effects due to the substrate and excited states (see Methods).

Although graphene was oxidized and strained by underlying oxidized copper, it showed no cracks (in optical micrographs or SEM images) after transfer (Supplementary Figs 11 and 12). The sheet resistance of transferred graphene also recovered to 426 ohms per square ($\Omega\,\square^{-1}$) after heat treatment at 600 °C in vacuum for 4 h, which was comparable to the 400 $\Omega\,\square^{-1}$ of the pristine graphene (PMMA removed; see Methods for details). The recovery of the electrical properties of graphene is often useful for studies of GGB-related intrinsic characteristics of graphene.

To provide proof of concept and to demonstrate the power of our approach, several graphene samples were prepared using different growth temperatures during CVD. For each sample, the Cu foil was annealed for 30 min, followed by 5 min of graphene growth at a given temperature. The range of graphene grain sizes was estimated to be of the order of micrometres on the basis of TEM examination (Supplementary Fig. 13). The samples then received an ultraviolet treatment time identical to that described in the experimental section in Methods (10 min). We found that the above-mentioned black spots and GGB lines occurred regardless of the growth temperature (Fig. 4a–d). The number of graphene grains in an area 57 μm × 76 μm was counted to obtain an average grain size per unit area. As the growth temperature increased, the sheet resistance of the graphene gradually decreased to ~300 $\Omega\,\square^{-1}$ for a growth temperature of 1,060 °C, and the average graphene grain size increased to ~72 μm^2 (Fig. 4e). The copper grain size fluctuated within the range of 10,000–15,000 μm^2; however, a correlation between copper grain size and graphene growth temperature was not observed. The sheet resistance of the graphene decreased as its grain size increased. Other growth conditions with different annealing temperatures (while maintaining the same growth temperatures) were also evaluated but resulted in similar behaviours (Supplementary Fig. 14). Thus, increasing the graphene grain size is a key factor in improving the graphene quality, whereas changing the Cu grain size is not effective in altering the sheet resistance of graphene.

Conductance AFM measurements showed that different GGBs have different conductances, and these GGBs were found to be consistent with the morphological GGBs observed after ultraviolet exposure (Fig. 4g–i and Supplementary Figs 15 and 16). The sheet resistance can be fitted to $\Omega = \Omega_o[1 + (A/A_c)^{-n}]$, where Ω_o is the sheet resistance of the intrinsic graphene without grain boundary scattering (or an infinite grain size), A is the average grain size, A_c is a fitting parameter, and n is an exponent. We obtained a value of Ω_o (230 $\Omega\,\square^{-1}$), which to our knowledge has not been estimated previously. The theoretical limit of Ω_o is 30 $\Omega\,\square^{-1}$ (refs 23, 24). In our case, the value is larger owing to scattering from defects and the substrate within the grain. This implies that minimizing defects—such as point defects, wrinkles and ripples—and substrate scattering will be important for improving conductivity, in addition to enlarging the graphene grain size.

To demonstrate our method in a more advanced application, we visualized fracture propagation on bending, using optical microscopy (Fig. 4j–m). Fractures propagated preferentially normal to the strain direction and terminated at GGBs (white squares; Fig. 4k). As the radius of curvature decreased, more cracks were created and propagated through GGBs, and in some cases, propagation directions were altered at the GGB lines (Fig. 4l, m). Our method is not limited to graphene, and can be generalized to analyse the defects and grain size

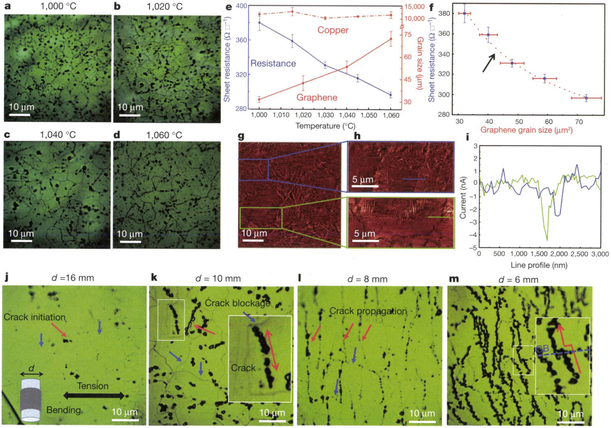


Figure 4 | The correlation between GGBs and sheet resistance. **a–d**, Oxidized graphene/Cu grown at indicated temperatures from 1,000 to 1,060 °C. **e**, **f**, Graphene sheet resistance versus graphene grain size for different growth conditions. Error bars, s.d. ($n = 4$). In **f**, a fitted curve for the sheet resistance is arrowed; $\Omega = \Omega_0[1 + (A/A_c)^{-n}]$ (see text for details). **g–i**, Conductance AFM of graphene/Cu before oxidation, showing two different types of GGBs. In **i**, line scans across small (big) grain boundaries are indicated by blue (green) curves. **j–m**, Fracture propagation in graphene on Cu as a function of radius of curvature ($d/2$, see inset in **j**). Black dots that appeared after oxidation show the position of cracks (red arrows). GGBs are indicated by blue arrows. In **k** and **m**, the left inset is shown magnified in the right inset. GB, grain boundary. Fractures propagated normal to the strain direction, terminated (**k**, **l**) and changed direction at the GGBs (**m**).

distribution of other two-dimensional layered structures, such as boron nitride (Supplementary Figs 17 and 18) and exfoliated clays.

METHODS SUMMARY

Ultraviolet oxidization of graphene on copper. Graphene on copper foil was placed into a chamber equipped with a low-pressure Hg lamp (LH-arc, Lichtzen, with an output of 20 mW cm^{-2}, with the majority of emitted light at a wavelength of 254 nm and approximately 10% of light at a wavelength of 185 nm)[13]. Humidity was introduced into the chamber by connecting it to a water bubbler. The humidity level in the chamber was monitored using a hygro-thermometer (accuracy of ±3%). The chamber was continuously ventilated throughout the experiments to maintain a constant pressure. After reaching the required humidity level, the water bubbler was disconnected from the chamber. The graphene/Cu substrate was then irradiated with ultraviolet light for 10 min to oxidize the samples. Humid air undergoes the following reactions under ultraviolet irradiation[21,22]: reactions (1), (2) and (4) require ultraviolet light to proceed, reaction (3) is thermally activated.

$$O_2 \rightarrow O_3 \tag{1}$$

$$O_3 + H_2O \rightarrow O_2 + H_2O_2 \tag{2}$$

$$2O_3 + H_2O_2 \rightarrow 2OH^{\bullet} + 3O_2 \tag{3}$$

$$H_2O \rightarrow H^{\bullet} + OH^{\bullet} \tag{4}$$

Full Methods and any associated references are available in the online version of the paper.

Received 23 May; accepted 4 September 2012.

Published online 3 October 2012.

1. Yazyev, O. V. & Louie, S. G. Electronic transport in polycrystalline graphene. *Nature Mater.* **9,** 806–809 (2010).
2. Yu, Q. *et al.* Control and characterization of individual grains and grain boundaries in graphene grown by chemical vapour deposition. *Nature Mater.* **10,** 443–449 (2011).
3. Huang, P. Y. *et al.* Grains and grain boundaries in single-layer graphene atomic patchwork quilts. *Nature* **469,** 389–392 (2011).
4. Kim, K. *et al.* Grain boundary mapping in polycrystalline graphene. *ACS Nano* **5,** 2142–2146 (2011).
5. Tian, J., Cao, H., Wu, W., Yu, Q. & Chen, Y. P. Direct imaging of graphene edges: atomic structure and electronic scattering. *Nano Lett.* **11,** 3663–3668 (2011).
6. Rasool, H. I. *et al.* Atomic-scale characterization of graphene grown on copper (100) single crystals. *J. Am. Chem. Soc.* **133,** 12536–12543 (2011).
7. Rasool, H. I. *et al.* Continuity of graphene on polycrystalline copper. *Nano Lett.* **11,** 251–256 (2011).
8. Gao, L., Guest, J. R. & Guisinger, N. P. Epitaxial graphene on Cu(111). *Nano Lett.* **10,** 3512–3516 (2010).
9. Kim, D. W., Kim, Y. H., Jeong, H. S. & Jung, H.-T. Direct visualization of large-area graphene domains and boundaries by optical birefringency. *Nature Nanotechnol.* **7,** 29–34 (2012).
10. Blake, P. *et al.* Making graphene visible. *Appl. Phys. Lett.* **91,** 063124 (2007).
11. Li, X. *et al.* Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **324,** 1312–1314 (2009).
12. Han, G. H. *et al.* Influence of copper morphology in forming nucleation seeds for graphene growth. *Nano Lett.* **11,** 4144–4148 (2011).
13. Güneş, F. *et al.* UV-light-assisted oxidative sp^3 hybridization of graphene. *NANO* **6,** 409–418 (2011).
14. Jin, Z. *et al.* Click chemistry on solution-dispersed graphene and monolayer CVD graphene. *Chem. Mater.* **23,** 3362–3370 (2011).
15. Malarda, L. M. Pimentaa, M. A., Dresselhaus, G. & Dresselhaus, M. S. Raman spectroscopy in graphene. *Phys. Rep.* **473,** 51–87 (2009).
16. Huang, M. *et al.* Phonon softening and crystallographic orientation of strained graphene studied by Raman spectroscopy. *Proc. Natl Acad. Sci.* **106,** 7304–7308 (2009).
17. Chae, S. J. *et al.* Synthesis of large-area graphene layers on poly-nickel substrate by chemical vapor deposition: wrinkle formation. *Adv. Mater.* **21,** 2328–2333 (2009).
18. Liu, Y. & Yakobson, B. I. Cones, pringles, and grain boundary landscapes in graphene topology. *Nano Lett.* **10,** 2178–2183 (2010).
19. Lee, G., Lee, B., Kim, J. & Cho, K. Ozone adsorption on graphene: ab initio study and experimental validation. *J. Phys. Chem. C* **113,** 14225–14229 (2009).
20. Jandhyala, S. *et al.* Atomic layer deposition of dielectrics on graphene using reversibly physisorbed ozone. *ACS Nano* **6,** 2722–2730 (2012).
21. Feiyan, C., Pehkonen, S. O. & Ray, M. B. Kinetics and mechanisms of UV-photodegradation of chlorinated organics in the gas phase. *Wat. Res.* **36,** 4203–4214 (2002).
22. Wang, J. H. & Ray, M. B. Application of ultraviolet photooxidation to remove organic pollutants in the gas phase. *Separ. Purif. Tech.* **19,** 11–20 (2000).
23. Chen, J.-H., Jang, C., Xiao, S., Ishigami, M. & Fuhrer, M. S. Intrinsic and extrinsic performance limits of graphene devices on SiO$_2$. *Nature Nanotechnol.* **3,** 206–209 (2008).
24. Jeong, C., Nair, P., Khan, M., Lundstrom, M. & Alam, M. A. Prospects for nanowire-doped polycrystalline graphene films for ultratransparent, highly conductive electrodes. *Nano Lett.* **11,** 5020–5025 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the Star Faculty programme (2010-0029653), the International Research and Development programme (2011-00242) and the WCU programme (R31-2008-10029) of the NRF of Korea funded by MEST.

Author Contributions D.L.D. and G.H.H. contributed equally to this work in experiment planning, experiment measurements, data analysis and manuscript preparation. S.M.L. performed the theoretical calculations. F.G. prepared the samples for TEM measurements. The copper grain size was characterized by H.K. SEM and EDS were performed by E.S.K. and J.W.J. S.H.C. and S.C.L. designed the ultraviolet chamber and humidity controller. S.T.K. and S.C.L. performed conductance AFM. K.P.S. designed and performed the bending test. Graphene on nickel was prepared by S.J.C. S.J.Y. performed recovery of sheet resistance after ultraviolet exposure. Q.H.T. prepared the graphene samples for all the experiments. The TEM images were taken by M.H.P. S.M.L. and J.Y.C. contributed to the manuscript preparation. Y.H.L. contributed to experiment planning, data analysis and manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.H.L. (leeyoung@skku.edu).

doi:10.1038/nature11520

April 2012 intra-oceanic seismicity off Sumatra boosted by the Banda-Aceh megathrust

Matthias Delescluse[1], Nicolas Chamot-Rooke[1], Rodolphe Cattin[2], Luce Fleitout[1], Olga Trubienko[1] & Christophe Vigny[1]

Large earthquakes nucleate at tectonic plate boundaries, and their occurrence within a plate's interior remains rare and poorly documented, especially offshore. The two large earthquakes that struck the northeastern Indian Ocean on 11 April 2012 are an exception: they are the largest strike-slip events reported in historical times[1,2] and triggered large aftershocks worldwide[3]. Yet they occurred within an intra-oceanic setting along the fossil fabric of the extinct Wharton basin, rather than on a discrete plate boundary[4–8]. Here we show that the 11 April 2012 twin earthquakes are part of a continuing boost of the intraplate deformation between India and Australia that followed the Aceh 2004 and Nias 2005 megathrust earthquakes, subsequent to a stress transfer process recognized at other subduction zones[9,10]. Using Coulomb stress change calculations, we show that the coseismic slips of the Aceh and Nias earthquakes can promote oceanic left-lateral strike-slip earthquakes on pre-existing meridian-aligned fault planes. We further show that persistent viscous relaxation in the asthenospheric mantle several years after the Aceh megathrust explains the time lag between the 2004 megathrust and the 2012 intraplate events. On a short timescale, the 2012 events provide new evidence for the interplay between megathrusts at the subduction interface and intraplate deformation offshore. On a longer geological timescale, the Australian plate, driven by slab-pull forces at the Sunda trench, is detaching from the Indian plate, which is subjected to resisting forces at the Himalayan front[6,8,11].

The first and largest shock of 11 April 2012 (moment magnitude M_{w} 8.6) occurred 330 km west-southwest of the Aceh 2004 epicentre, but only 120 km southwest of the subduction front. It was followed 2 h later on the same day by an M_{w} 8.2 twin earthquake, located some 180 km south of the main event. The oceanic fabric there is characterized by a set of more or less north–south-oriented fracture zones, well captured by satellite-derived gravity measurements (Fig. 1) and mapped *in situ* using multi-beam bathymetry[4]. These are fossil transform faults that were active during the opening of the Wharton basin that shut off some 45 Myr ago. Recent reactivation of these faults[12] was confirmed by seismic reflection profiling showing basement and sediment offsets as well as sea-floor expression[13]. The epicentres of the two 11 April 2012 events align along a 022° direction, which is within a few degrees of one of the nodal planes for both earthquakes.

The two twin shocks thus seem to have involved slip on one of these very long fossil fracture zones that can be followed for several hundred kilometres[6]. The perpendicular normal fault orientation only offers limited continuous fault length to the rupture as it is offset by fracture zones[14]. However, we cannot rule out slip on the conjugate plane[15] as such slip happened further south in the Wharton basin in the year 2000[7]. For the M_{w} 8.6 event, aftershocks (Fig. 1) and back-projection analysis indicate the activation of N110°E planes[16,17]. The entire oceanic fabric, normal faults included, seems to be somewhat reactivated by these complex events.

The 11 April 2012 earthquake sequence is located in an area previously recognized as part of the diffuse deformation zone between the Indian and Australian plates[18]. Intraplate deformation is active on

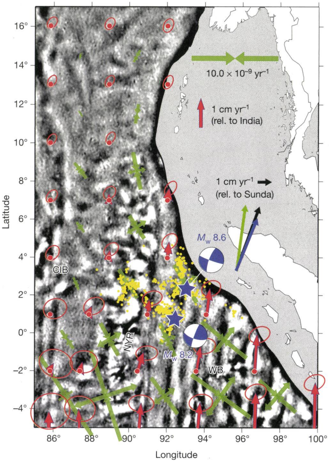


Figure 1 | Present-day kinematics of the India–Australia plate. The 11 April 2012 M_{w} 8.6 and M_{w} 8.2 earthquakes, shown here by their blue and white focal mechanisms ('beachballs'), occurred in the Wharton basin (WB), at the heart of the diffuse plate boundary between India and Australia. Ninetyeast Ridge (NYR), which is the Kerguelen hotspot trail, separates the Wharton basin from the central Indian basin (CIB). Off trench, the background is the 120 km high-pass-filtered satellite free-air gravity anomaly, illuminating meridian aligned fracture zones. Three convergence vectors are displayed east of the Sumatra-Andaman trench, calculated at the location of the 2004 Aceh epicentre. The blue vector is India/Sunda vector (IN/SU) predicted by the MORVEL[33] global model (46 mm yr^{-1} towards N22°E), while the green vector is the Australia/Sunda prediction using the same model (54 mm yr^{-1} towards N8°E). The true convergence vector is in between, since the subducting plate is neither India nor Australia. The black vector (56 mm yr^{-1} towards N20°E) is an estimation of the convergence rate that takes into account a non-rigid India-Australia plate. This non-rigid crustal velocity field is shown as red vectors with their 95% error ellipses at knot points west of the trench line. The corresponding strain-rate field is shown as green double arrows. At the latitude of the April 2012 earthquake sequence, motion of the Wharton basin with respect to India progressively increases west to east, from negligible close to Ninetyeast ridge to 10 mm yr^{-1} close to the trench.

---

[1] Laboratoire de Géologie, Ecole Normale Supérieure and CNRS UMR8538, 75005 Paris, France. [2] Géosciences Montpellier, Université Montpellier II and CNRS UMR5243, 34090 Montpellier, France.

both sides of Ninetyeast Ridge, predominantly as reverse faults with north–south compressive axes (P-axes) west of it[19–21] and strike-slip faults with northwest–southeast P-axes east of it[4,8]. Large-scale folding of the oceanic lithosphere was found both in the equatorial Indian Ocean and in the Wharton basin, and interpreted as a direct consequence of large in-plate stresses[11]. The present-day kinematics of the Indian and Australian plates is now well established, based on Global Positioning System (GPS) measurement away from the deforming zone. Following the approach of Holt and Haines[5,22] (Methods), we have derived a self-consistent non-rigid velocity field[8] (Fig. 1) that combines far-field boundary conditions (GPS over Australia and India, plus a few islands within the deforming zone) and local constraints from the deformation zone itself (style of deformation from focal mechanisms and heat flow as a proxy for strength). Offshore from Sumatra, the motion exhibits a gradual south–north evolution from the Australian plate velocity to the Indian plate velocity. The predicted amount of shear across the set of meridian aligned faults is of the order of $10\,\mathrm{mm\,yr^{-1}}$.

Intraplate seismicity has been boosted by the Aceh and Nias ruptures. We show in Fig. 2 a chrono-spatial chart of the seismicity of the Wharton basin, covering the 8 years before and the 8 years after the Aceh earthquake. An increase in the rate of seismicity immediately followed the Aceh and Nias earthquakes, especially for the portions of oceanic lithosphere experiencing peak slip rate on the failure plane of the Aceh megathrust. The Nias earthquake that followed the Aceh earthquake triggered oceanic earthquakes in a region that remained silent during the time lag of three months between the two events. We also estimated the net increase in intraplate seismic moment release. On the basis of a century-long catalogue[8], the mean moment release for the entire deforming Indian Ocean (including the central Indian basin and the Wharton basin) was about $2.3 \times 10^{19}\,\mathrm{N\,m\,yr^{-1}}$ before the Aceh earthquake, two-thirds of it being released in the Wharton basin. In the year 2005, this number jumped to $1.2 \times 10^{20}\,\mathrm{N\,m}$ for the Wharton basin only. For comparison, seismic moment release for the two 11 April earthquakes is above $1.4 \times 10^{22}\,\mathrm{N\,m}$, which is about 1,000 times the yearly release estimated from the century-long catalogue.

We compiled focal mechanisms of 47 earthquakes ($M_w > 5$) that occurred off Sumatra in the oceanic plate from December 2004 (Aceh earthquake) until 15 April 2012 (from the Global CMT catalogue). We excluded earthquakes within the slab portions already engaged in subduction, although many of them are clearly related to intraplate deformation and reactivation of the oceanic fabric[7], such as the Padang earthquake of 2009 (M_w 7.6, depth >80 km). The post-Aceh oceanic earthquakes fall into two categories: normal faults at the external wall of the trench (18 events), and earthquakes with a northwest–southeast P-axis corresponding to the release of India-Australia intraplate stress (29 earthquakes). Among the compressive events, only one seems unrelated to the oceanic fabric, the other 28 being strike-slip earthquakes with one left-lateral meridian aligned nodal plane (015° ± 005°). The very consistent orientation of the nodal planes over a wide region is strong evidence for reactivation of the oceanic fabric in the area.

To test the effect of the Aceh 2004 and Nias 2005 ruptures on the reactivation of the Wharton basin oceanic fabric, we calculated the far-field static Coulomb stress change $\Delta\mathrm{CFF} = \Delta\tau - \mu'\Delta\sigma_n$ resulting from the coseismic displacement on the subduction interface, up to 250 km seaward of the trench axis. Here $\Delta\tau$ is the static shear stress change on the failure planes (positive in the direction of fault slip), $\Delta\sigma_n$ the static normal stress change (positive if the fault is clamped) and μ' the effective friction coefficient (see Methods and ref. 23). A positive Coulomb stress change favours rupture.

Calculations were first performed to resolve stress onto a N15°E-oriented vertical plane (Fig. 3a). The Coulomb stress change is calculated at 18 km depth, which is the average depth of the earthquakes before 2012 and an upper limit for the centroid depth of the mainshock of 11 April 2012 (18–27 km; ref. 14). We find that all strike-slip earthquakes occurred in a lobe of positive Coulomb stress change. The area of highest Coulomb stress change in front of the 2004 Aceh coseismic

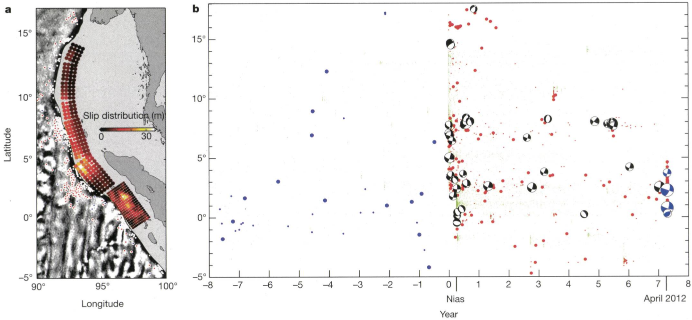


Figure 2 | Oceanic intraplate seismicity boost after the Aceh (2004) and Nias (2005) earthquakes. a, Map representation; **b**, chronological chart. **a**, Map representation of the same intraplate seismicity as **b**. To ensure the intraplate nature of the earthquakes, events 50 km off-trench are shown as larger dots. **b**, Year 0 corresponds to 26 December 2004. Red (after December 2004) and blue (before December 2004) events are exclusively located over the oceanic plate west of the trench line and up to 250 km away from the trench line, whereas green points represent the entire seismicity, including the upper plate. All earthquakes were extracted from the NEIC catalogue. Focal mechanisms of post-Aceh events with magnitudes greater than 5 occurred after December 2004 are also displayed. Blue focal mechanisms correspond to the April 2012 earthquake sequence, and black focal mechanisms to earthquakes that occurred in the same area since the Aceh event.

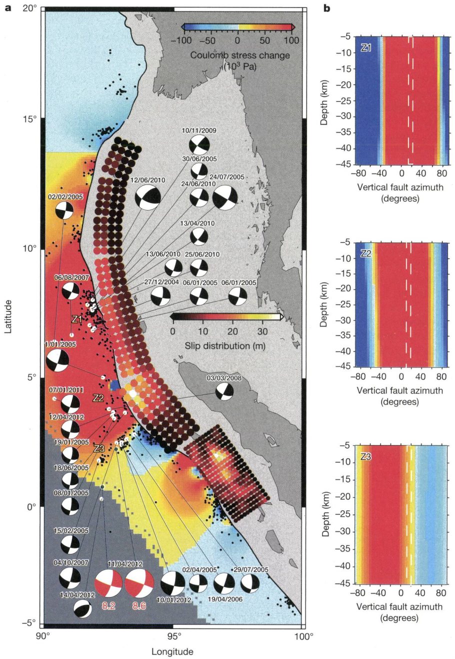

a
b
20°
15°
10°
5°
0°
−5°
90°
95°
100°
Latitude
Longitude
−100 −50 0 50 100
Coulomb stress change (10³ Pa)
0 10 20 30
Slip distribution (m)
10/11/2009
30/06/2005
24/07/2005
12/06/2010
24/06/2010
02/02/2005
13/04/2010
13/06/2010
25/06/2010
06/08/2007
27/12/2004
06/01/2005
06/01/2005
Z1
1/01/2005
03/03/2008
07/01/2011
Z2
12/04/2012
19/01/2005
Z3
18/06/2005
08/01/2005
15/02/2005
04/10/2007
11/04/2012
02/04/2005
29/07/2005
14/04/2012
8.2
8.6
10/01/2012
19/04/2006
Z1
Z2
Z3
Depth (km)
−5 −10 −15 −20 −25 −30 −35 −40 −45
−80 −40 0 40 80
Vertical fault azimuth (degrees)

Figure 3 | Static Coulomb stress change calculation offshore the Sumatra-Andaman subduction. a, Each colour pixel represents a N15°E-oriented vertical fault plane. Note that the slip colour scale is adapted to the Aceh earthquake slip distribution, and requires division by a factor of three to obtain the Nias earthquake slip values. **b,** The Coulomb stress change has been tested for different depths and azimuths of the fault planes at three different locations, from north to south: Z1 (30 June 2005 earthquake location), Z2 (1 January 2005 earthquake location) and Z3 (11 April 2012 mainshock location). White bars bound the azimuths of the oceanic fabric. Slip on the oceanic fabric is best favoured to the north of the area.

slip corresponds to a shear stress increase and to a normal stress decrease. This situation favours left-lateral strike-slip on the N15°E vertical faults independently of the value of the effective friction coefficient (unclamping). We further tested the effect of the azimuth of the vertical planes and the variation of the Coulomb stress change with depth (Fig. 3b). The mainshock of 11 April 2012 locates the exact limit where N20°E planes are still positively favoured by the Aceh 2004 slip. The Coulomb stress change is much higher if we consider that the M_w 8.6 event nucleated on a N110°E plane with dextral motion (see Supplementary Information), as suggested by seismological studies[16,17]. Depth does not really affect the result for the considered 5–45 km range. The 2005 Nias rupture also favours the occurrence of earthquakes on N15°E-oriented vertical planes, but focal mechanisms there indicate normal rather than strike-slip events. Most of the normal faults oriented along the trench strike are located in areas of positive Coulomb stress change for their respective receiver fault orientation (see Supplementary Information). Slip on newly formed thrust faults perpendicular to the intraplate P-axis are inhibited by the Aceh earthquake coseismic stresses (see Supplementary Information). Over the long term, only a few reverse faults were active during the past century[7], as stresses apparently rarely reached the level required to create a new fault.

The Coulomb stress change distribution is highly sensitive to the coseismic slip distribution. The high sensitivity of the model to shallow slip, in particular, may partly be due to assumptions underlying the Coulomb model, such as a flat-plane subduction interface and an homogeneous elastic half-space. Many coseismic solutions (see, for example, refs 24–29) have been proposed for the Aceh and Nias earthquakes based on the inversion of various data sets, including GPS displacements, coral reef observations, long-period teleseismic data, tide-gauge records and satellite altimetry measurements. We use here the slip distributions obtained in refs 24 and 25, for the Aceh and Nias earthquakes, respectively. First-order properties of the slip distributions obtained in these two models are consistent with several other studies[26–29]. Our best choice for slip distribution was however driven by one significant observation: triggered intraplate earthquakes, both normal (see Supplementary Information) and strike-slip, occur extremely close to the toe of the trench. Our sensitivity tests showed that only slip distributions including shallow, near trench displacement successfully produce a positive Coulomb stress change close enough to the trench[9]. Most coseismic slip inversions are rather insensitive to shallow displacements at the toe of the subduction, which is generally too far from the land stations where measurements are available. They thus tend to minimize the amount of shallow rupture there. However, marine constraints[30] (2011 Tohoku M_w 9.0 earthquake, Japan), and geodetic measurements close to the failure plane[31] (2010 Maule M_w 8.8 earthquake, Chile) have shown that the coseismic displacement of giant earthquakes does reach the toe of the trench.

A clear drawback of the Coulomb stress calculation is that it fails to model time-dependent stress transfers because the underlying rheology is elastic. In these elastic models, promoted faults should break shortly after the megathrust event, as the extensional stress within the oceanic plate off the trench begins to decrease immediately after the earthquake with the start of a new interseismic loading cycle. Yet the time lag between Aceh and the 2012 earthquake sequence is more than 7 years, which indicates another mechanism producing a northeast–southwest unclamping that is larger than the slow buildup of northeast–southwest compressive stress during interseismic loading. Transient triggering may be related to the viscoelastic properties of the asthenosphere. In Supplementary Information, we present a simple two-dimensional viscoelastic model of the earthquake cycle that predicts the time evolution of the stresses linked to subduction earthquakes as a function of distance to the trench. The observed time lag between the Aceh and Nias earthquakes and the April 2012 earthquake sequence is found to be compatible with a rheology of the asthenosphere independently required to explain near field to far field GPS constraints in the Sunda block, on the other side of the trench[32]. The maximum post-seismic relaxation stress is reached between 7 and 10 years after the megathrust in the area of the 2012 events. The further from the trench, the longer the time lag until maximum post-seismic relaxation stress is reached. Similar three-dimensional finite element viscoelastic calculations[32], using a more realistic rupture zone geometry, further confirm that at the locations of the two events of April 2012, the net stress effect is unclamping in the northeast direction, in agreement with the two-dimensional viscoelastic model and the Coulomb stress change calculations presented here.

The two large earthquakes of 11 April 2012 are typical of the intraplate deformation between the Indian and Australian plates, the fossil fabric of the oceanic lithosphere being re-activated across the entire Indian Ocean[19,21]. The mega-earthquakes of Aceh and Nias acted as instantaneous boosters, suddenly illuminating where and how intraplate deformation was at work. Viscous relaxation in the asthenospheric mantle is responsible for short-term stress building, still contributing several years to several tens of years after the megathrust events. Long-term stress building there[11] is the direct consequence of the high mechanical coupling of the Indian plate to the Eurasian plate at the Himalayan front that started some 8 Myr ago. Dense oceanic slabs engaged at the Sunda trench inexorably drive the Australian plate northward, while this motion is resisted at the Himalayas. The long-term scenario is that a nascent plate tectonic boundary is forming: the Australian plate is becoming detached from the Indian plate.

METHODS SUMMARY

Kinematic modelling. The kinematic model is detailed elsewhere[8]. It includes constraints from far-field GPS measurements and style of deformation from a century-long intraplate earthquake catalogue.

Coulomb stress change calculations. Coseismic stresses are computed using megathrust slip distributions described as displacement discontinuities in an homogeneous elastic half-space[23]. The shear and normal stresses are then extracted for a chosen plane geometry that is supposed to pre-exist in the entire oceanic lithosphere.

Post-seismic relaxation. A two-dimensional finite element model featuring variable rheological parameters is used (Z-set/Zébulon; see Methods). The rheological parameters fit the upper plate GPS constraints.

Full Methods and any associated references are available in the online version of the paper.

Received 30 April; accepted 23 August 2012.

Published online 26 September 2012.

1. Robinson, D. P. A rare great earthquake on an oceanic fossil fracture zone. *Geophys. J. Int.* **186,** 1121–1134 (2011).
2. McGuire, J. J. & Beroza, G. C. A rogue earthquake off Sumatra. *Science* **336,** 1118–1119 (2012).
3. Pollitz, F. F., Stein, R. S., Sevilgen, V. & Bürgmann, R. The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide. *Nature* http://dx.doi.org/10.1038/nature11504 (this issue).
4. Deplus, C. *et al.* Direct evidence of active deformation in the eastern Indian oceanic plate. *Geology* **26,** 131–134 (1998).
5. Tinnon, M., Holt, W. E. & Haines, A. J. Velocity gradients in the northern Indian Ocean inferred from earthquake moment tensors and relative plate velocities. *J. Geophys. Res.* **100,** 24315–24329 (1995).
6. Deplus, C. Indian Ocean actively deforms. *Science* **292,** 1850–1851 (2001).
7. Abercrombie, R. E., Antolik, M. & Ekstrom, G. The June 2000 Mw 7.9 earthquakes south of Sumatra: deformation in the India-Australia plate. *J. Geophys. Res.* **108,** 2018, http://dx.doi.org/10.1029/2001JB000674 (2003).
8. Delescluse, M. & Chamot-Rooke, N. Instantaneous deformation and kinematics of the India-Australia plate. *Geophys. J. Int.* **168,** 818–842 (2007).
9. Dmowska, R., Rice, J. R., Lovison, L. C. & Josell, D. Stress transfer and seismic phenomena in coupled subduction zones during the earthquake cycle. *J. Geophys. Res.* **93,** 7869–7884 (1988).
10. Lay, T., Ammon, C. J., Kanamori, H., Xue, L. & Kim, M. J. Possible large near-trench slip during the 2011 Mw 9.0 off the Pacific coast Tohoku earthquake. *Earth Planets Space* **63,** 713–718 (2011).
11. Gerbault, M. At what stress level is the central Indian Ocean lithosphere buckling? *Earth Planet. Sci. Lett.* **178,** 165–181 (2000).
12. Rajendran, K., Andrade, V. & Rajendran, C. P. The June 2010 Nicobar earthquake: fault reactivation on the subducting oceanic plate. *Bull. Seismol. Soc. Am.* **101,** 2568–2577 (2011).
13. Graindorge, D. *et al.* Impact of lower plate structure on upper plate deformation at the NW Sumatran convergent margin from seafloor morphology. *Earth Planet. Sci. Lett.* **275,** 201–210 (2008).
14. Satriano, C., Kiraly, E., Bernard, P. & Vilotte, J. P. The 2012 Mw 8.6 Sumatra earthquake: evidence of westward sequential seismic ruptures associated to the reactivation of a N-S ocean fabric. *Geophys. Res. Lett.* **39**, L00B07, http://dx.doi.org/10.1029/2012GL052387 (2012).
15. Hwang, L. J. & Kanamori, H. Rupture process of the 1987–1988 Gulf of Alaska earthquake sequence. *J. Geophys. Res.* **97,** 19881–19908 (1992).
16. Yue, H., Lay, T. & Koper, K. D. *En échelon* and orthogonal fault ruptures of the 11 April 2012 great intraplate earthquakes. *Nature* http://dx.doi.org/10.1038/nature11492 (this issue).
17. Meng, L. *et al.* Earthquake in a maze: compressional rupture branching during the 2012 M_w 8.6 Sumatra earthquake. *Science* **337,** 724–726 (2012).
18. Royer, J. Y., Gordon, R. G., DeMets, C. & Vogt, P. R. New limits on the motion between India and Australia since Chron 5 (11 Ma) and implications for lithospheric deformation in the equatorial Indian Ocean. *Geophys. J. Int.* **129,** 41–74 (1997).
19. Bull, J. M. & Scrutton, R. A. Fault reactivation in the central Indian Ocean and the rheology of oceanic lithosphere. *Nature* **344,** 855–858 (1990).
20. Chamot-Rooke, N. *et al.* Intraplate shortening in the central Indian Ocean determined from a 2100-km-long north-south deep seismic reflection profile. *Geology* **21,** 1043–1046 (1993).
21. Delescluse, M., Montési, L. G. J. & Chamot-Rooke, N. Fault reactivation and selective abandonment in the oceanic lithosphere. *Geophys. Res. Lett.* **35,** L16312, http://dx.doi.org/10.1029/2008GL035066 (2008).
22. Haines, A. J. & Holt, W. E. A procedure for obtaining the complete horizontal motions within zones of distributed deformation from the inversion of strain rate data. *J. Geophys. Res.* **98,** 12057–12082 (1993).
23. Cattin, R. *et al.* Stress change and effective friction coefficient along the Sumatra-Andaman-Sagaing fault system after the 26 December 2004 (Mw=9.2) and the 28 March 2005 (Mw=8.7) earthquakes. *Geochem. Geophys. Geosyst.* **10,** Q03011, http://dx.doi.org/10.1029/2008GC002167 (2009).
24. Rhie, J., Dreger, D., Bürgmann, R. & Romanowicz, B. Slip of the 2004 Sumatra-Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets. *Bull. Seismol. Soc. Am.* **97,** S115–S127 (2007).

25. Ji, C. Preliminary result of the March 28, 2005 Mw 8.68 Nias earthquake. http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2005/03/smooth/nias.html (2005).
26. Vigny, C. *et al.* Insight into the 2004 Sumatra-Andaman earthquake from GPS measurements in southeast Asia. *Nature* **436,** 201–206 (2005).
27. Chlieh, M. *et al.* Coseismic slip and afterslip of the great Mw 9.15 Sumatra-Andaman earthquake of 2004. *Bull. Seismol. Soc. Am.* **97,** S152–S173 (2007).
28. Briggs, R. *et al.* Deformation and slip along the Sunda megathrust in the great 2005 Nias-Simeulue earthquake. *Science* **311,** 1897–1901 (2006).
29. Banerjee, P., Pollitz, F., Nagarajan, B. & Bürgmann, R. Coseismic slip distributions of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Nias earthquakes from GPS static offsets. *Bull. Seismol. Soc. Am.* **97,** S86–S102 (2007).
30. Fujiwara, T. *et al.* The 2011 Tohoku-Oki earthquake: displacement reaching the trench axis. *Science* **334,** 1240 (2011).
31. Vigny, C. *et al.* The 2010 Mw 8.8 Maule megathrust earthquake of Central Chile, monitored by GPS. *Science* **332,** 1417–1421 (2011).
32. Fleitout, L. *et al.* Far away motions associated with giant subduction earthquakes and the mechanical properties of the lithosphere-asthenosphere system. *Geophys. Res. Abs.*, **14,** EGU2012–10899 (2012 EGU General Assembly, 2012).
33. DeMets, C., Gordon, R. G. & Argus, D. F. Geologically current plate motions. *Geophys. J. Int.* **181,** 1–80 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements Epicentres of intraplate earthquakes were obtained from the National Earthquake Information Center (NEIC) catalogue. Focal mechanisms were obtained from the Global Centroid-Moment-Tensor (CMT) Project. Figures were prepared with GMT software. We thank R. Abercrombie for comments.

Author Contributions M.D. and N.C.-R. wrote the manuscript and prepared most of the figures with contributions from all co-authors. R.C. wrote the Coulomb stress code, L.F. and O.T. prepared the post-seismic relaxation figures and texts. C.V. has been measuring GPS velocities in southeast Asia for years, some of them being used in the kinematic model.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.D. (delescluse@geologie.ens.fr).

LETTER

doi:10.1038/nature11492

En échelon and orthogonal fault ruptures of the 11 April 2012 great intraplate earthquakes

Han Yue[1], Thorne Lay[1] & Keith D. Koper[2]

The Indo-Australian plate is undergoing distributed internal deformation caused by the lateral transition along its northern boundary—from an environment of continental collision to an island arc subduction zone[1,2]. On 11 April 2012, one of the largest strike-slip earthquakes ever recorded (seismic moment magnitude M_w 8.7) occurred about 100–200 kilometres southwest of the Sumatra subduction zone. Occurrence of great intraplate strike-slip faulting located seaward of a subduction zone is unusual. It results from northwest–southeast compression within the plate caused by the India–Eurasia continental collision to the northwest, together with northeast–southwest extension associated with slab pull stresses as the plate underthrusts Sumatra to the northeast. Here we use seismic wave analyses to reveal that the 11 April 2012 event had an extraordinarily complex four-fault rupture lasting about 160 seconds, and was followed approximately two hours later by a great (M_w 8.2) aftershock. The mainshock rupture initially expanded bilaterally with large slip (20–30 metres) on a right-lateral strike-slip fault trending west-northwest to east-southeast (WNW–ESE), and then bilateral rupture was triggered on an orthogonal left-lateral strike-slip fault trending north-northeast to south-southwest (NNE–SSW) that crosses the first fault. This was followed by westward rupture on a second WNW–ESE strike-slip fault offset about 150 kilometres towards the southwest from the first fault. Finally, rupture was triggered on another *en échelon* WNW–ESE fault about 330 kilometres west of the epicentre crossing the Ninetyeast ridge. The great aftershock, with an epicentre located 185 kilometres to the SSW of the mainshock epicentre, ruptured bilaterally on a NNE–SSW fault. The complex faulting limits our resolution of the slip distribution. These great ruptures on a lattice of strike-slip faults that extend through the crust and a further 30–40 kilometres into the upper mantle represent large lithospheric deformation that may eventually lead to a localized boundary between the Indian and Australian plates.

It has long been recognized from relative plate motions that the Indo-Australian plate is not behaving as a single rigid unit[1]. Earthquake focal mechanisms, plate spreading rates inferred from magnetic lineations, fracture zone orientations, seismic stratigraphy, folds and sedimentary unconformities, and geodetic observations indicate diffuse internal deformation of the plate over a broad equatorial region (Fig. 1). This region extends from the central Indian ridge near the Chagos bank, eastward past the Ninetyeast ridge to the Sumatra trench[2–5], southward along the Ninetyeast ridge[6], and southeastward throughout the Wharton basin[5,7]. The southwestern part of the plate appears to have already fragmented to produce the Capricorn subplate[5,8], which has a diffuse border with the Australian plate along the southern Ninetyeast ridge. The NNE–SSW trends of the Ninetyeast ridge and fracture zones in the Wharton basin, together with aligned left-lateral strike-slip faulting mechanisms in both areas, may lead one to anticipate a similar orientation for any great rupture in the intraplate deformation zone. For example, the large 18 June 2000 (13.87° S, 97.3° E; M_w 7.9) earthquake in the Wharton basin appears to have involved predominantly left-lateral strike-slip faulting along the expected NNE–SSW orientation, although a second fault orientation was also activated during the 34-s-long rupture[9,10]. However, the east–west trend of the equatorial deformation zone raises the possibility of right-lateral faulting further north.

On 11 April 2012, a great intraplate earthquake (M_w 8.7) initiated at 2.31° N, 93.06° E at 08:38:37 UTC, followed by a great aftershock (M_w 8.2) at 0.77° N, 92.45° E at 10:43:09 UTC (Fig. 1)[11]. The overall faulting geometries of both events inferred from point-source moment

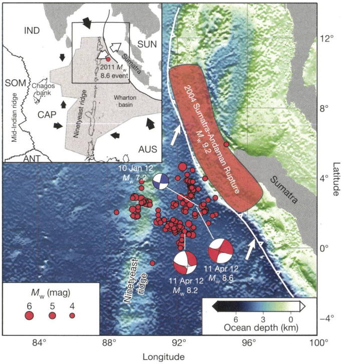


Figure 1 | The 11 April 2012 rupture sequence. Top inset, the regional plate tectonic setting, with the Indo-Australian plate being segmented into three subplates—India (IND), Australian (AUS) and Capricorn (CAP); also shown are the adjacent Somalian (SOM), Antarctic (ANT) and Sunda (SUN) plates. The light grey zone is a region of intraplate deformation between the subplates. Black arrows indicate directions of intraplate compression and white arrows show extension from subduction[5]. Major bathymetric features like the Chagos bank, Ninetyeast ridge and Wharton basin are indicated. Main panel, the USGS W-phase moment tensor solutions for the great events of 11 April 2012 (beachballs), USGS one-week aftershock locations (red circles with magnitude scale at lower left), and location and focal mechanism for the 10 January 2012 (M_w 7.2) event used as empirical Green functions in the surface wave analysis. The barbed line indicates the Sumatra trench, and the rupture zone of the M_w 9.2 Sumatra-Andaman underthrusting event[15] is shown. Bathymetry is shown, with the colour scale at the lower right. White arrows indicate the direction of motion of the Indo-Australian composite plate relative to the Sunda plate.

---

[1] Department of Earth and Planetary Sciences, University of California Santa Cruz, Santa Cruz, California 95064, USA. [2] Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112, USA.

tensor representations are similar (Fig. 1), with steeply dipping strike-slip orientations having either left-lateral slip on NNE–SSW faults or right-lateral slip on WNW–ESE faults, either of which would be consistent with the pervasive northwest–southeast compressional stress orientation throughout the region[7]. The first event is possibly the largest strike-slip earthquake ever seismically recorded (the 15 August 1950 Assam (M_{w} 8.6) interplate earthquake in the eastern Himalayas is of comparable size[12], but uncertain in faulting mechanism[13–15]), and it is probably the largest intraplate earthquake ever seismically recorded. This event triggered increases in seismicity globally[16]. The 2012 events are located 100–200 km seaward of the Sumatra subduction zone in which the Indo-Australian plate is underthrusting the Sunda plate, offshore of the epicentral region of the 26 December 2004 Sumatra-Andaman (M_{w} 9.2) interplate earthquake rupture (Fig. 1). The 2004 event involved 15–30 m of trench-perpendicular coseismic displacement on the plate boundary[17], and increased northeast horizontal extensional stress in the 2012 source region[18]. The ocean lithosphere varies in age from about 45 to 65 million years from east to west across the source region, although thermal rejuvenation along the Ninetyeast ridge may reduce the effective age there by up to 20 million years (ref. 19).

The relative epicentral locations of the two great events on 11 April 2012 immediately suggest rupture on a NNE–SSW-striking fault, and numerous rapid analyses of the seismic waves performed to characterize the space-time faulting history made this assumption; however, as more aftershock locations were determined, clear trends along parallel WNW–ESE lineations offset in latitude by ~150 km soon became apparent (Fig. 1). A robust seismic method for identifying fault planes and rupture spatial extent is by stacking short-period P-wave signals from networks of seismic stations at teleseismic distances corrected for propagation from a grid of possible source locations[20–23]. Within a few hours of the events, such semi-automated back-projections of short-period P waves from the mainshock recorded by networks of stations in Europe, China, Japan and Australia showed that pulses of coherent short-period seismic energy release appeared to illuminate both WNW aftershock trends, as well as additional loci of seismic radiation that were clearly not on a single NNE–SSW-oriented fault[24].

A summary of our own back-projections of short-period teleseismic P waves recorded by stations around Europe is presented in Fig. 2 (a time-varying animation is provided in Supplementary Movie 1). Coherent sources of short-period seismic energy radiation are imaged for more than 160 s and display a close correspondence with the regional distribution of epicentres of early (first-week) aftershocks, which are clearly on multiple faults. The initial rupture on the northeastern WNW–ESE fault appears to expand bilaterally, with stronger radiation in the western part of the fault. After about 50 s the WNW rupture propagation ends, with a total fault length of about 150 km and a low overall rupture velocity of 1.5–2.0 $\mathrm{km\,s^{-1}}$. Around 40 s into the rupture process, seismic radiation begins to be emitted from a perpendicular, presumably left-lateral conjugate plane that first ruptures 50–100 km in the SSW direction (from 30 to 60 s), then 50–100 km in the NNE direction (from 55 to 90 s). After about 70 s, seismic radiation continues on a second WNW–ESE fault that is roughly parallel to the first and separated by ~150 km to the southwest. This rupture propagates to the WNW, perhaps discontinuously, until about 145 s, at which time seismic radiation is apparent on a fourth distinct fault further to the west and persists until ~160 s. Also shown in Fig. 2 are back-projection results for the M_{w} 8.2 aftershock that commences just over 2 h later. The region of short-period energy release is much more compact than that of the mainshock, and the duration is roughly half as long (~80–85 s). The short-period energy release indicates bilateral rupture on a NNE–SSW plane with dominant propagation initially from the epicentre towards the NNE, consistent with the locations of early aftershocks, and weaker late energy release to the SSW.

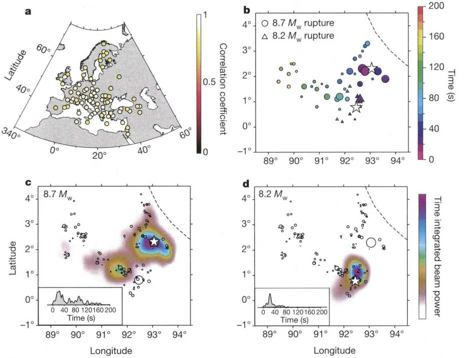


Figure 2 | Short-period seismic energy release pattern. a, Locations of seismograph stations in Europe that were used for the 0.5–2.0 Hz back-projection of the M_{w} 8.7 and M_{w} 8.2 Sumatra earthquakes. Signals were selected for high P-wave similarity (P-wave signal correlation coefficients relative to the array average are indicated) and a broad geographical distribution. **b,** Local beam power maxima during the back-projections, with colour indicating time after the nominal origin time. Symbol size is proportional to beam power. The dashed line is the location of the Sumatra trench. White stars are the epicentral locations of the M_{w} 8.7 and 8.2 earthquakes. **c,** Normalized time integrated relative beam power in units of $\mathrm{cm^2\,s^{-2}}$ (colour scale at right ranges linearly from white for zero to purple for unity) for the M_{w} 8.7 back-projection. The circles are early NEIC aftershocks with symbol size proportional to magnitude, and the dashed line is the Sumatra trench. The inset shows peak beam power as a function of time. **d,** Same as **c** but for the M_{w} 8.2 aftershock. The time-dependent behaviour is best evaluated by viewing Supplementary Movies 1 and 2.

Similar patterns of short-period radiation are observed for arrays of stations in Japan (see Supplementary Fig. 1 and Supplementary Movies 1 and 2), and by other investigators who have posted preliminary solutions online[24–27]. Using a lower-frequency passband of 0.1–0.5 Hz for the European P-wave observations smears the energy more broadly in space and accentuates the late energy in the mainshock rupture that occurs near longitudes of 89.5°–90.5° E, but otherwise gives results very similar to those in Fig. 2 (Supplementary Movies 1 and 2).

An important limitation of these and other short-period back-projections is that they are not directly sensitive to magnitude of fault slip (short-period seismic radiation is sensitive to slip-rate variations, and back-projection methods depend critically on wave field coherency, which can be high for spatially concentrated radiation from a small event and low for diffuse radiation from a large event)[22,23,28]. Thus, whereas the short-period results indicate multiple potential activated faults, the relative seismic moments of the fault segments are not resolved.

To overcome this deficiency, we used broadband surface waves to image spatial variations in seismic moment release, again allowing for the possibility of multiple fault segments. To account for complex surface wave propagation effects, we use global seismic recordings of short-arc Rayleigh waves (R_1) and Love waves (G_1) for a moderate size (M_w 7.2) event on 10 January 2012 with a source location (2.45° N, 93.21° E; 18:36:59 UTC) about 20 km northeast of the mainshock epicentre (Fig. 1), and a similar strike-slip focal mechanism. These signals are deconvolved from the corresponding M_w 8.7 and 8.2 great-event recordings to produce time series called source time functions[29] (STFs). The large earthquake rupture properties can be inferred by projecting the STFs into a space-time history of seismic radiation in essentially the same way as is done in short-period back-projections.

The spatial distributions of long-period seismic wave radiation imaged over a gridded region around the M_w 8.7 and 8.2 earthquake epicentres are shown in Fig. 3. Darker colours indicate stronger long-period source radiation, which tracks the aftershock distribution. There is some intrinsic smearing of the images due to the non-uniform and incomplete azimuthal coverage, and the spatial resolution is lower than for the short-period back-projections because we are dealing with one-sided moment rate functions and lower frequency signals; however, this method provides a better indication of seismic moment distribution. Plots of this imaging with R_1 and G_1 separately are shown in Supplementary Fig. 2.

For the M_w 8.7 event, the largest long-period source energy is concentrated near the epicentre, but there is significant source energy along both ESE–WNW aftershock trends, indicating either continuous rupture on corresponding faults or a sequence of discrete coseismically triggered 'early' aftershocks. The secondary features have peak amplitudes from 20–30% of the largest features, indicative of significant relative seismic moment. Snapshots of the reconstructed long-period radiation as a function of time are shown in Supplementary Fig. 3; these reveal rupture propagation in the WNW direction on both faults, and that the concentration of radiation almost 400 km west of the epicentre occurs at about 120 s. The relative STFs for the M_w 8.2 aftershock were similarly processed (Supplementary Fig. 4), and indicate source radiation concentrated near the hypocentre with rupture propagation towards the NNE. Comparison with Fig. 2 shows good agreement in map locations of source radiation for ~1 s period energy

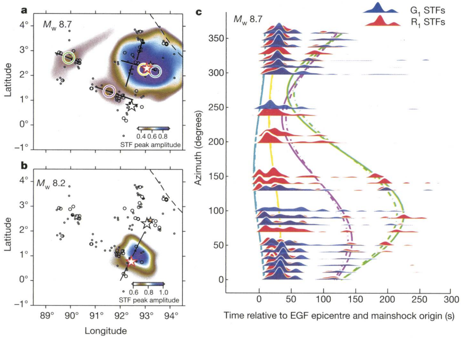


Figure 3 | Long-period seismic energy release pattern. **a**, **b**, Maps showing the spatial patterns of long-period surface wave energy sources for the 11 April 2012 (M_w 8.7) mainshock (**a**; epicentre indicated with a red outlined star) and the large (M_w 8.2) aftershock 2 h later (**b**; epicentre indicated with a red outlined star). The one-week aftershock distribution is shown by small circles with radii proportional to seismic magnitude. The colour images indicate the peak stacked energy at each position from combined R_1 source time functions (STFs) and G_1 STFs back-projected relative to the epicentral location of the 10 January 2012 (M_w 7.2) event (orange stars) which was used for empirical Green functions (EGFs). Solid lines indicate the orientations of likely subfaults that ruptured during each event. Snapshots that indicate the time history of energy release are shown in Supplementary Fig. 3. Coloured circles in the mainshock image indicate specific locations of energy release that produce arrivals considered in **c**, which shows the azimuthally binned and stacked R_1 (red) and G_1 (blue) STFs plotted in time relative to the EGF epicentre. These represent seismic moment as a function of time observed at different azimuths. The coloured curves correspond to the coloured circles in **a**, showing the predicted arrival times of energy from each location, with 'S-curve' patterns due to relative propagation times towards different azimuths. The solid curves are for R_1 waves using a velocity of 4.0 km s^{-1} and the dashed curves are for G_1 waves using a velocity of 4.5 km s^{-1}. The features aligned on such curves form the corresponding images in **a**, although images formed after 60 s of delay from the STF onsets have the first 60 s of the STFs masked out to avoid contamination. A corresponding profile of STFs for the M_w 8.2 aftershock is shown in Supplementary Fig. 4.

and ~20–200 s period energy, bolstering the interpretation that multiple faults with significant slip were activated during the main rupture. Neither imaging method provides depth resolution.

To resolve the spatial slip distribution, broadband teleseismic P and SH waves were inverted together with the surface wave STFs for finite-faulting models using multiple distinct fault segments inferred from the back-projections of short- and long-period seismic energy. Although single-fault inversions allow basic waveform features in the first 60 s to be fairly well modelled, the overall waveforms can be fitted better with additional fault complexity (and more parameters). Finite-fault inversions require specification of substantial a priori information about faulting geometry, rupture initiation time, rupture velocity, and discretization of the fault surface and subfault STFs. We use the consistency of the short-period and long-period imaging approaches in conjunction with the aftershock trends to specify four primary subfaults for the mainshock and one segment for the largest aftershock, with locations and timing of rupture initiation guided by the high-resolution short-period back-projections. Dip δ and rake λ orientations for each subfault with strike ϕ were constrained by extensive modelling and inversion suites.

Figure 4 shows the fault model geometry and the slip distributions obtained by least-squares inversion of P and SH waves using up to 200-s-long time windows of the seismograms (with some being truncated early to avoid contamination from PP and SS phases) along with the R_1 and G_1 STFs. Guided by the short-period back-projections, we specified the rupture velocity as $2.0\ \mathrm{km\,s^{-1}}$ on all segments. There is uncertainty in the kinematic parameters, but the collective information from aftershock locations and P wave and surface wave back-projections does provide first-order constraints.

Our results show that the great events of 11 April 2012 involve rupture of a very complex network of faults, for which we have no documented precedent in recorded seismic history. Good fits are found for the large P and SH wave and R_1 and G_1 data set (see Supplementary Fig. 5), although some secondary features are not fully accounted for. The failure process can be summarized as follows: first, a large moment release during a predominantly WNW-propagating, asymmetric bilateral strike-slip rupture (with $2.0\ \mathrm{km\,s^{-1}}$ velocity) with large peak slip (~37 m) and about 150 km total rupture length with seismic moment corresponding to M_w ~8.5. Second, this rupture triggered bilateral failure of a cross-cutting orthogonal fault that had a seismic moment corresponding to M_w ~7.9. Third, subsequent rupture occurred to the south on an *en échelon* ESE–WNW fault that expanded in the WNW direction with a seismic moment of M_w ~8.3. Fourth, the process culminated with a fourth activated fault segment about 300 km west of the epicentre that ruptured on either a WNW–ESE or NNE–SSW fault (or both) with M_w ~7.8 (Supplementary Fig. 6 shows the inversion results assuming the NNE–SSW orientation). The cumulative moment of these ruptures is 13.9×10^{21} N m, which gives M_w 8.7 (about 40% larger than point-source seismic moment estimates, and 15% larger than a two-subevent seismic moment estimate[30]). A modest non-double couple component is found when the individual subfault moment tensors are summed directly or allowing for temporal shifts (Supplementary Fig. 7).

Even allowing for the trade-offs and non-uniqueness of the very complex model description, it is well-established that this event activated a complex lattice of faults in the deformation zone between the India and Australia subplates, with the deep centroid depths, large estimated fault displacements and large extent of faulting of the sequence suggesting localization of deformation in the region extending westward from Sumatra to the Ninetyeast ridge. The shortening between the India and Australia subplates that is being accommodated across this deformation zone is mainly being distributed onto strike-slip faults rather than thrust faults, and the ultimate configuration of the plate boundary that will develop is difficult to anticipate. The failure process is somewhat influenced by the plate fabric, with NNE–SSW-trending structures embedded in the plate from its earlier formation probably providing zones of weakness that can fail in strike-slip events. However, the primary seismic moment is on WNW–ESE features, which cross-cut the trend of the Ninetyeast ridge. Large rupture through oceanic lithosphere cross-cutting fracture zones has been observed in the Antarctic plate[31]. High-resolution bathymetry swaths along the ridge have revealed many WNW–ESE fractures in the seafloor before the 2012 sequence[32], and these young features appear to

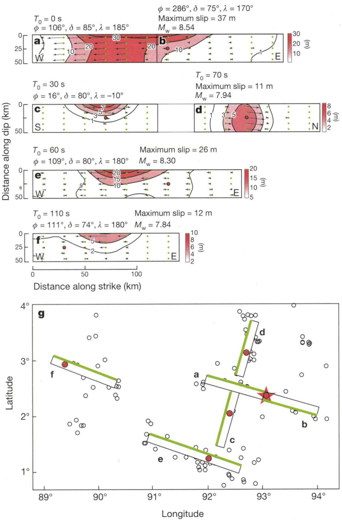


Figure 4 | Map of primary faulting during the M_w 8.7 event. a–f, Slip distributions on each subfault; **g**, map showing locations of subfaults. Map shows fault segments activated during the 11 April 2012 M_w 8.7 mainshock, specified for finite-fault slip model inversion of teleseismic broadband P and SH waves and R_1 and G_1 STFs. The faulting complexity is guided by the short-period and long-period source imaging in Figs 2 and 3 (and associated animations and time snapshots in the Supplementary Information) along with the one-week aftershock distribution. The red star indicates the epicentral location, and red dots indicate the placement of hypocentres (all at 30 km depth) on each fault segment. The rectangles indicate the subfault strikes and dips (shallow edge is on the green side, deeper edge on the black side. We use four subfaults (two of the subfaults are further subdivided) with onset times (T_0) constrained by the short-period back-projections. Slip distributions on each subfault (arrows indicate the relative size and direction of slip, with slip size colour contoured using the colour scale at right; green dots indicate the grid) are shown in **a–f**, with the subfault grids having 20 km spacing along-strike and 10 km spacing along-dip. The peak slip and M_w of each subfault is indicated, as is the position of the hypocentre on each subfault from which the rupture expands. The rupture velocity is $2.0\ \mathrm{km\,s^{-1}}$ for all subfaults. The M_w 8.2 event (epicentre given by black star) is not inverted because the broadband P and SH waves are obscured from surface waves from the first event, but it appears to involve bilateral rupture on a fifth fault, trending NNE–SSW. δ, dip; λ, rake; ϕ, strike.

have dominated the faulting in this great earthquake sequence. Many more large fracturing events will be needed to evolve a localized plate boundary, so future extraordinary ruptures in the region may well occur, but this event currently stands as the largest-magnitude intraplate strike-slip earthquake geophysicists have yet recorded.

METHODS SUMMARY

For a gridded distribution of possible rupture locations across the source region, short-period body waves and long-period surface waves from teleseismic stations were back-projected to image the space-time patterns of seismic wave radiation for the M_w 8.7 and 8.2 events, with minimal a priori assumptions about faulting geometry. Narrow-band filtered P waves from large aperture arrays in Europe and Japan were used to image locations of short-period radiation on the source grid using predicted travel-time variations across the arrays relative to an initial hypocentral alignment by shifting and fourth-root summing of the signals. Broadband R_1 Rayleigh waves and G_1 Love waves from global seismic stations were used to image the space-time locations of long-period seismic radiation. The complex dispersion effects of surface wave propagation were first removed using corresponding recordings from an M_w 7.2 earthquake on 10 January 2012 with very similar location and focal mechanism to the M_w 8.7 mainshock as empirical Green functions (EGFs). The EGF signals were deconvolved from the records of the large events, eliminating long-range propagation effects and extracting relative source time functions (STFs) for the large event. The STFs were then back-projected onto the source grid using average phase velocities for R_1 (4.0 km s^{-1}) and for G_1 (4.5 km s^{-1}). We then specified a set of four fault segments composed of multiple 10-km-wide by 20-km-long subfaults that activate at prescribed times (based on the short-period imaging), and invert broadband teleseismic P and SH waves and the surface wave STFs for finite fault slip distributions. The geometries of the segments are constrained by aftershock distribution and focal mechanisms, back-projections, and suites of inversions with varying parameters. The final model has time-varying slip on each fault segment consistent with the back-projection imaging, aftershocks, and broadband teleseismic signals.

Full Methods and any associated references are available in the online version of the paper.

Received 6 May; accepted 7 August 2012.
Published online 26 September 2012.

1. Minster, J. B. & Jordan, T. H. Present day plate motions. *J. Geophys. Res.* **83,** 5331–5354 (1978).
2. Wiens, D. *et al.* A diffuse plate boundary model for Indian Ocean tectonics. *Geophys. Res. Lett.* **12,** 429–432 (1985).
3. Gordon, R. G., DeMets, C. & Argus, D. F. Kinematic constraints on distributed lithospheric deformation in the equatorial Indian Ocean from present motions between the Australian and Indian plates. *Tectonics* **9,** 409–422 (1990).
4. Gordon, R. G. DeMets, C. & Royer, J.-Y. Evidence for long-term diffuse deformation of the lithosphere of the equatorial Indian Ocean. *Nature* **395,** 370–374 (1998).
5. Royer, J.-Y. & Gordon, R. G. The motion and boundary between the Capricorn and Australian plates. *Science* **277,** 1268–1274 (1997).
6. Stein, S. & Okal, E. A. Seismicity and tectonics of the Ninetyeast Ridge area, evidence for internal deformation of the Indian Plate. *J. Geophys. Res.* **83,** 2233–2245 (1978).
7. Delescluse, M. & Chamot-Rooke, N. Instantaneous deformation and kinematics of the India-Australia plate. *Geophys. J. Int.* **168,** 818–842 (2007).
8. DeMets, C. & Royer, J.-Y. A new high-resolution model for India-Capricorn motion since 20 Ma: implications for the chronology and magnitude of distributed crustal deformation in the Central Indian Basin. *Curr. Sci.* **85,** 339–345 (2003).
9. Robinson, D. P., Henry, C., Das, S. & Woodhouse, H. H. Simultaneous rupture along two conjugate planes of the Wharton Basin earthquake. *Science* **292,** 1145–1148 (2001).
10. Abercrombie, R. E., Antolik, M. & Ekström, G. The June 2000 M_w 7.9 earthquakes south of Sumatra: deformation in the India-Australia Plate. *J. Geophys. Res.* **108** (B1), 2018, http://dx.doi.org/10.1029/2001JB000674 (2003).
11. McGuire, J. J. & Beroza, G. C. A rogue earthquake off Sumatra. *Science* **336,** 1118–1119 (2012).
12. Kanamori, H. The energy release in great earthquakes. *J. Geophys. Res.* **82,** 2981–2987 (1977).
13. Ben-Menahem, A., Aboodi, E. & Schild, R. The source of the great Assam earthquake — an interplate wedge motion. *Phys. Earth Planet. Inter.* **9,** 265–289 (1974).
14. Chen, W.-P. & Molnar, P. Seismic moments of major earthquakes and the average rate of slip in Central Eurasia. *J. Geophys. Res.* **82,** 2945–2969 (1977).
15. Molnar, P. A review of the seismicity and the rates of active underthrusting and deformation at the Himalayas. *J. Himal. Geol.* **1,** 131–154 (1990).
16. Pollitz, F. F., Stein, R. S., Sevilgen, V. & Bürgmann, R. The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide. *Nature* http://dx.doi.org/nature11504 (this issue).
17. Ammon, C. J. *et al.* Rupture process of the 2004 Sumatra-Andaman earthquake. *Science* **308,** 1133–1139 (2005).
18. Delescluse, M. *et al.* April 2012 intra-oceanic seismicity off Sumatra boosted by the Banda–Aceh megathrust. *Nature* http://dx.doi.org/nature11520 (this issue).
19. Shapiro, N. M., Ritzwoller, M. H. & Engdahl, E. R. Structural context of the great Sumatra-Andaman Islands earthquake. *Geophys. Res. Lett.* **35,** L05301, http://dx.doi.org/10.1029/2008GL033381 (2008).
20. Ishii, M., Shearer, P. M., Houston, H. & Vidale, J. E. Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-net array. *Nature* **435,** 933–936 (2005).
21. Krüger, F. & Ohrnberger, M. Tracking the rupture of the M_w 9.3 Sumatra earthquake over 1,150 km at teleseismic distance. *Nature* **435,** 937–939 (2005).
22. Koper, K. D., Hutko, A. R., Lay, T. & Sufri, O. Imaging short-period seismic radiation from the 27 February 2010 Chile (M_w 8.8) earthquake by back-projection of P, PP, and PKIKP waves. *J. Geophys. Res.* **117,** B02308, http://dx.doi.org/10.1029/2011JB008576 (2012).
23. Koper, K. D., Hutko, A. R. & Lay, T. Along-dip variation of teleseismic short-period radiation from the 11 March 2011 Tohoku Earthquake (M_w 9.0). *Geophys. Res. Lett.* **38,** L21309, http://dx.doi.org/10.1029/2011GL049689 (2011).
24. Incorporated Research Institutions for Seismology. Back projections for M_w 8. 7 off W coast of Northern Sumatra, http://www.iris.edu/spud/backprojection/118733 (2012).
25. Meng, L., Ampuero, J.-P. & Luo, Y. Back-projection results, 4/11/2012 (Mw8.6) offshore Sumatra, Indonesia. http://www.tectonics.caltech.edu/slip_history/2012_Sumatra/back_projection/.
26. Kiser, E. Preliminary rupture modeling of the April 11, 2012 Sumatran earthquakes. http://www.seismology.harvard.edu/research_sumatra2012.html.
27. Wang, D., Mori, J. & Ohmi, S. Rupture process of the April 11, 2012 Sumatra (Mw 8.6) earthquake imaged with back-projection of Hi-net data. http://www.eqh.dpri.kyoto-u.ac.jp/src/etc/sumatra.htm.
28. Lay, T. *et al.* Depth-varying rupture properties of subduction zone megathrust faults. *J. Geophys. Res.* **117,** B04311, http://dx.doi.org/10.1029/2011JB009133 (2012).
29. Lay, T. *et al.* The 2006–2007 Kuril Islands great earthquake sequence. *J. Geophys. Res.* **114,** B11308, http://dx.doi.org/10.1029/2008JB006280 (2009).
30. Duputel, Z. *et al.* The 2012 Sumatra great earthquake sequence. *Earth Planet. Sci. Lett.* **351-352,** http://dx.doi.org/10.1016/j.epsl.2012.07.017 (2012).
31. Hjörleifsdóttir, V., Kanamori, H. & Tromp, J. Modeling 3-D wave propagation and finite slip for the 1998 Balleny Islands earthquake. *J. Geophys. Res.* **114,** B03301, http://dx.doi.org/10.1029/2008JB005975 (2009).
32. Meng, L. *et al.* Earthquake in a maze: compressional rupture branching during the 2012 M_w 8.6 Sumatra earthquake. *Science* **337,** 724–726 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank H. Kanamori, Z. Duputel and G. Hayes for discussions and exchanges of information about this event. A. Hutko provided early short-period back-projection results. We thank R. Abercrombie for comments on this paper. This work made use of GMT and SAC software and Federation of Digital Seismic Networks (FDSN) seismic data. The Incorporated Research Institutions for Seismology (IRIS) Data Management System (DMS), the European ORFEUS Data Center and the NIED F-net Data Centre were used to access the data. This work was supported by NSF grant EAR0635570 (T.L.) and EAR0951558 (K.D.K.).

Author Contributions H.Y. contributed to the surface wave back-projections and finite fault modelling; K.D.K. performed the short-period back-projections; and T.L. performed finite-fault inversions and guided the synthesis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.L. (tlay@ucsc.edu).

LETTER

doi:10.1038/nature11504

The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide

Fred F. Pollitz[1], Ross S. Stein[1], Volkan Sevilgen[2] & Roland Bürgmann[3]

Large earthquakes trigger very small earthquakes globally during passage of the seismic waves and during the following several hours to days[1–10], but so far remote aftershocks of moment magnitude $M \geq 5.5$ have not been identified[11], with the lone exception of an $M = 6.9$ quake remotely triggered by the surface waves from an $M = 6.6$ quake 4,800 kilometres away[12]. The 2012 east Indian Ocean earthquake that had a moment magnitude of 8.6 is the largest strike-slip event ever recorded. Here we show that the rate of occurrence of remote $M \geq 5.5$ earthquakes (>1,500 kilometres from the epicentre) increased nearly fivefold for six days after the 2012 event, and extended in magnitude to $M \leq 7$. These global aftershocks were located along the four lobes of Love-wave radiation; all struck where the dynamic shear strain is calculated to exceed 10^{-7} for at least 100 seconds during dynamic-wave passage. The other $M \geq 8.5$ mainshocks during the past decade are thrusts; after these events, the global rate of occurrence of remote $M \geq 5.5$ events increased by about one-third the rate following the 2012 shock and lasted for only two days, a weaker but possibly real increase. We suggest that the unprecedented delayed triggering power of the 2012 earthquake may have arisen because of its strike-slip source geometry or because the event struck at a time of an unusually low global earthquake rate, perhaps increasing the number of nucleation sites that were very close to failure.

The 11 April 2012 $M = 8.6$ east Indian Ocean earthquake (Fig. 1a) is by far the largest strike-slip event ever recorded[13]. It was a complex event rupturing a series of subparallel and conjugate faults with dominant moment release within a ~100-s time span[14–19]. This intra-oceanic earthquake precipitated a large, abrupt increase in seismicity worldwide (Fig. 2a). This seismicity reached remote distances of 10,000–20,000 km from the mainshock (Fig. 1b, c), much wider than the near-field aftershock zone surrounding a large mainshock, which is generally a few fault lengths in size[20]. The seismicity jump involved increases in earthquake productivity across a broad magnitude spectrum (Fig. 3). At rates well above background, several $M \geq 5.5$ aftershocks were triggered, unlike those in all previous remote-triggering cases[1–10].

The global distribution of these remote aftershocks is consistent with Love-wave radiation from the mainshock. We consider the square root of the second invariant of the deviatoric strain tensor, ε^{II}, the dynamic shear strain realized during passage of the seismic waves, using a point source convolved with a 100-s-long source time function that replicates observed seismic waveforms (Methods Summary).

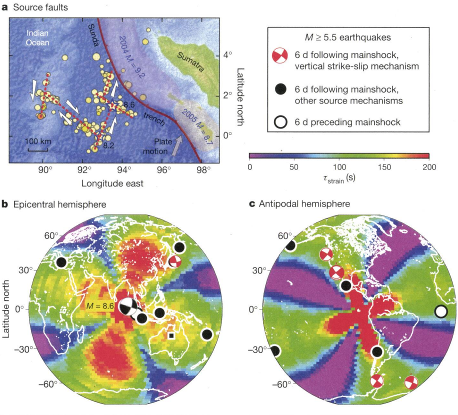


Figure 1 | The 2012 $M = 8.6$ mainshock and $M = 8.2$ aftershock fault ruptures and maps of strain duration τ_{strain} at a threshold value of 0.1 microstrain. a, Inferred fault ruptures of the 11 April 2012 $M = 8.6$ east Indian Ocean earthquake and an $M = 8.2$ aftershock that occurred 2 h later. Superimposed are the first 20 d of $M \geq 4.5$ aftershocks of 0–100-km depth. These earthquakes probably ruptured a complex set of subparallel and conjugate faults with the indicated sense of motion[14–19] (arrows). Parts of the rupture areas of the 2004 $M = 9.2$ and 2005 $M = 8.7$ Nias earthquakes on the Sunda megathrust are indicated. **b, c,** Global maps of τ_{strain} (colour scale). Superimposed are the epicentres of $M \geq 5.5$ events that occurred during the 6 d preceding the mainshock (2 epicentres) and following the mainshock (24 epicentres, 16 of which are remote, that is, >1,500 km from the mainshock). Focal mechanisms of six post-mainshock events with near-vertical strike-slip mechanisms (plunge of neutral axis, >60°) are indicated with red beachballs. The 9:00:09 11 April 2012 $M = 5.5$ event (in the western Aleutian Islands) occurred 21 min 33 s after the mainshock between the direct P- and S-wave arrivals from the mainshock; all others are delayed by hours to days. The focal mechanism of the mainshock is plotted at its epicentre.

---

[1]US Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, California 94025, USA. [2]Seismicity.net, 490 Laurel Street, Suite 10, San Carlos, California 94070, USA. [3]Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA.

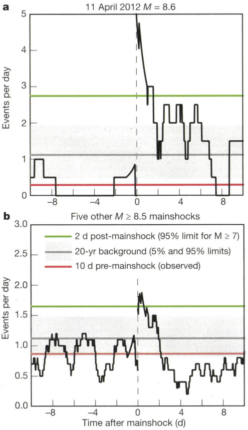


Figure 2 | Global rates of shallow (depth, ≤100 km) $M \geq 5.5$ earthquakes during the 10 d preceding and following a mainshock. Events within 1,500 km of the mainshock are excluded. A running average with a half-width of 1 d is used to construct each curve. Time is relative to the mainshock origin time. **a**, 11 April 2012 $M = 8.6$ east Indian Ocean mainshock. **b**, A collection of five other mainshocks with $M \geq 8.5$. Red horizontal lines denote the $M \geq 5.5$ seismicity rate for the 10 d preceding the 2012 mainshock (**a**) or event-averaged rates for the 10 d preceding the five other mainshocks (**b**). Gray horizontal lines and grey shading denote mean $M \geq 5.5$ seismicity rates and the 5% and 95% empirical probability bounds obtained from analysis of a 20-yr-long NEIC catalogue (Supplementary Fig. 4 and Methods section on background seismicity rates). Green horizontal lines give the 95% empirical upper bounds on R_{post} for one event (**a**) or an aggregate of five events (**b**) derived from a set of catalogue $M \geq 7$ mainshocks (Supplementary Fig. 7).

We define a first measure of dynamic strain, the 'strain duration' τ_{strain}, as the amount of time that ε^{II} exceeds a threshold value during the wave propagation. Figure 1 reveals that $M \geq 5.5$ aftershocks lie within the four lobes of high τ_{strain} (above a threshold strain of 10^{-7}, or 0.1 microstrain, for >100 s), which coincide with the theoretical lobes of maximum Love-wave excitation for a mainshock of its strike-slip source geometry. A second measure is the peak value of ε^{II} imparted by the mainshock during passage of the seismic waves. Maps of this peak (Supplementary Fig. 1) reveal that $M \geq 5.5$ aftershocks similarly lie within the four lobes of peak dynamic shear strain (>0.25 microstrain). Both measures are important for rationalizing the occurrence of dynamically triggered events[10,21,22].

These unprecedented observations raise the question of whether large aftershocks are always triggered at great distances by large mainshocks. Only very small dynamically triggered remote aftershocks of $M \geq 7$ mainshocks have been found[9], and $M > 5$ aftershocks during the first ~30 h have been found to be triggered only within ~1,000 km from $M \geq 7$ mainshocks[11] (generally within the range of static Coulomb stress triggering), suggesting that the global hazard of large aftershocks does not increase following even a large mainshock. Thus, the $M = 8.6$ east Indian Ocean triggering is probably quite rare.

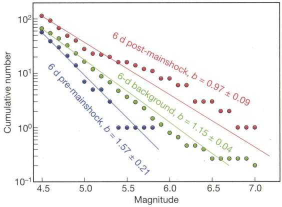


Figure 3 | Cumulative number of global $M \geq 4.5$ events of depth ≤100 km during the 6 d before and after the 2012 event. Both the pre-mainshock (blue symbols) and the post-mainshock (red symbols) seismicities are restricted to be remote (>1,500 km from the mainshock). Superimposed is the cumulative number of background 'remote' events in an average 6-d interval during the year preceding the 2012 mainshock (green symbols). Lines illustrate the corresponding b values from the Gutenberg–Richter law and standard deviations derived using maximum-likelihood regression.

To confirm that the global rate increase is real, we test the null hypothesis that the rate of remote $M \geq 5.5$ events is the same for periods before and after $M \geq 8.5$ mainshocks (Table 1), including the 2012 east Indian Ocean earthquake. The remoteness criterion is defined with a simple distance threshold from the mainshock centroid and is applied uniformly to both pre-mainshock and post-mainshock seismicity. Because the enhancement of dynamically triggered seismicity is thought to take place on a 1–2-d timescale[11], we test whether the average rate of remote $M \geq 5.5$ aftershocks in the 2 d following a mainshock, R_{post}, is the same as the rate, R_{pre}, of 'remote' $M \geq 5.5$ events in the 10 d preceding the mainshock. We use 2 d post-mainshock because of the observed duration of the rate increase and 10 d pre-mainshock to obtain a reliable background rate. Results are similar using other pre-mainshock windows, and here we address the statistical fluctuations in pre-mainshock rates to be expected from the chosen time window. We define the rate change to be

$$\Delta R = R_{\text{post}} - R_{\text{pre}} \qquad (1)$$

and test the null hypothesis by comparing the observed ΔR with empirical probability distributions of ΔR derived from Monte Carlo sampling of the 20-yr global US National Earthquake Information Center (NEIC) catalogue of $M \geq 5.5$ earthquakes (see Methods sections on magnitude of completeness and seismicity rate change).

There is a fivefold increase in the rate, R, of shallow global $M \geq 5.5$ events for several days following the 2012 mainshock (Fig. 2a). A similar but weaker rate change is observed for a collection of five other $M \geq 8.5$ mainshocks during the past decade (Fig. 2b and Table 1), exceeding the 10-d-average pre-mainshock rate by a factor of two for 2 d. The jump in seismicity rate is not attributable to induced mainshock–aftershock sequences (that is, aftershocks triggering aftershocks) and is also apparent over longer time intervals (Supplementary Fig. 2).

Table 1 | $M \geq 8.5$ earthquakes

Date (dd/mm/yyyy)	Magnitude*	Region/name	Tectonic environment	Sense of slip
26/12/2004	9.2†	Sumatra	Sunda megathrust	Thrust + strike slip
28/03/2005	8.7†	Nias	Sunda megathrust	Thrust
12/09/2007	8.5	Sumatra	Sunda megathrust	Thrust
27/02/2010	8.8	Maule, Chile	Andean trench	Thrust
11/03/2011	9.0	Tohoku, Japan	Japan trench	Thrust
11/04/2012	8.6	East Indian Ocean	Intraplate	Strike slip

* From NEIC catalogue unless otherwise noted.
† Ref. 32.

The rate change ΔR provides a test of the null hypothesis. As summarized in Methods, we consider the set of 243 shallow-focus $M \geq 7$ mainshocks that occurred during the 20 yr before the 2012 $M = 8.6$ event; the magnitude of completeness is 5.5 (Supplementary Fig. 3). For a single mainshock, the 95% upper bound on the empirical probability distribution of ΔR is 1.25 events per day. The observed rate increase for the 2012 $M = 8.6$ mainshock (Fig. 2a) is $\Delta R = 2.7$ events per day, which far exceeds the 95% empirical bound. An increase ΔR exceeding 2.5 events per day is not found for any $M \geq 7$ mainshock in the 1992–2012 catalogue (Supplementary Fig. 5). If the search is extended back to 1982, then a ΔR value exceeding 2.5 events per day occurs for only one out of 337 $M \geq 7$ mainshocks in the catalogue (the 5 April 1990 $M = 7.4$ normal-faulting Mariana Islands earthquake). If remote $M \geq 5.5$ events are assumed to occur with a Poissonian probability distribution, then one out of 337 mainshocks translates into a 0.3% probability of the 2012 rate increase occurring by chance.

For the aggregate of five other $M \geq 8.5$ mainshocks, the 95% bound on the empirical probability distribution of ΔR is 0.55 events per day (Supplementary Fig. 6). The observed rate increase (Fig. 2b) is $\Delta R = 0.63$ events per day, a weaker signal than the 2012 rate increase. Thus, the null hypothesis can be rejected at the 99.7% level for the 2012 $M = 8.6$ mainshock and at the 95% level for the aggregate of five other very large mainshocks.

The global nature of the delayed triggered seismicity following the east Indian Ocean earthquake and its association with the Love-wave radiation pattern raise a number of issues. The first is whether strike-slip earthquakes promote triggering of moderate global aftershocks more than thrust sources. The 2012 event is the only $M \geq 8.5$ mainshock with a predominantly strike-slip mechanism (Table 1). A strike-slip event will generate horizontally polarized shear-wave (SH- and Love-wave) energy in four dominant directions, along and perpendicular to the fault strike[23]. Love waves will preferentially stress near-vertical strike-slip faults. Six of the 16 (37%) remote aftershocks of the 2012 event were strike-slip events on near-vertical faults (Fig. 1), whereas the background rate of $M \geq 5.5$ strike-slip events is 24%; four of the strike-slip aftershocks were among the six that occurred during the first 2 d (67%). In contrast, low-angle thrust events generate predominantly P–SV-wave energy with smaller azimuthal variation of radiated energy[23]. For example, the 2004 $M = 9.2$ Sumatra and 2005 $M = 8.7$ Nias earthquakes occurred near the 2012 $M = 8.6$ event but produced smaller remote seismic displacements (Supplementary Fig. 10). Assuming that dynamic strain scales with displacement[24], this qualitatively suggests an enhanced triggering potential for the 2012 event.

A strike-slip source geometry, however, is of secondary importance to earthquake magnitude. We examined the next-largest strike-slip sources of the past 15 yr (the 23 December 2004 $M = 8.1$ Tasman Sea and 28 March 1998 $M = 8.1$ Antarctic earthquakes, both of which were about five times smaller than the 2012 $M = 8.6$ event), but we do not find any increase in remote $M \geq 5.5$ seismicity. Nor do we see (using the Centennial Earthquake Catalog[25]) remote $M \geq 7$ aftershocks following the 1905 $M = 8.4$ Mongolian earthquake doublet.

The triggered aftershocks are not preferentially located in the near-field, where dynamic strain magnitudes are high, but rather are distributed uniformly over the globe (Fig. 1 and Supplementary Figs 1 and 8a). The strain duration (Fig. 1) is more uniformly distributed globally than the peak strain amplitude (Supplementary Fig. 1). This suggests that a dynamic strain threshold enables triggering, as suggested previously[21], and reinforces the notion that dynamic triggering depends on not only the amplitude of transient dynamic strains but also the duration for which they are applied[8,22]. Small events triggered by passage of seismic waves from the 2011 Tohoku earthquake are also distributed globally, with no preference for the near-field region[10], suggesting that dynamic triggering, whether instantaneous or delayed, depends on more than dynamic strain amplitude.

The second issue is the nature of the mechanism behind the delayed triggering of large aftershocks. When compared with the near-instantaneous surface-wave triggering of small global aftershocks, the 2–6-d timescale of global large aftershock activity implies a longer nucleation process. Near-instantaneous triggering of undetected foreshocks or slow-slip transients could initiate a cascade process culminating in earthquakes[26]. Surface waves from large earthquakes have been found to trigger deep-seated non-volcanic tremor in several subduction zones and on the San Andreas fault[27]. This non-volcanic tremor is believed to be associated with slow slip.

The third issue is whether the large jump in global seismicity following the 2012 earthquake is related to the very low seismicity rate preceding it. The 2012 quake struck after 6–12 d of exceptionally low global seismicity (Fig. 4 and Methods section on low pre-earthquake seismicity rate). We suspect that because the dynamic stress is oscillatory and brief, only faults very close to failure can be triggered. If earthquake nucleation sites age, or are stressed towards failure, at a roughly constant rate, then when a period of such extremely low

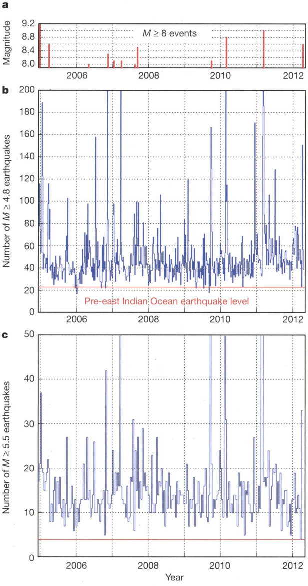


Figure 4 | Global seismicity rates during the 7.3 yr between the 2004 Sumatra earthquake and the 2012 east Indian Ocean earthquake. **a**, Magnitude of $M \geq 8$ events. **b**, **c**, Seismicity rates for $M \geq 4.8$ events in 6-d bins (**b**) and $M \geq 5.5$ events in 12-d bins (**c**). Catalogue is unedited; for example, no 1,500-km exclusion zones are used.

seismicity coincides with a great mainshock, there is a large reservoir of nucleation sites that are critically stressed or very close to failure. In contrast, none of the other $M \geq 8.5$ earthquakes is preceded by such a 6–12-d low background rate. Some of the close-to-failure sites may lie in aftershock zones of local mainshocks, such as one that is an aftershock of the 20 March 2012 $M = 7.4$ Mexican-trench event (Fig. 1).

The strongest evidence for delayed global triggering is provided by the preponderance of strike-slip post-mainshock events (Fig. 1), the jump in seismicity rate in individual regions (for example the Gulf of California (Supplementary Fig. 11)) and the observation that the $M \geq 5.5$ seismicity rate increase is part of a broader global $M \geq 4.5$ rate increase (Fig. 3), which involves large numbers of remote events (114 post-mainshock versus 57 pre-mainshock for 6-d periods).

The above considerations provide only partial answers to the question of why the 2012 event triggered so many remote large aftershocks. The apparent dependence of dynamic triggering on period, with longer periods having greater influence than shorter periods both experimentally and observationally[8,22,28], may be relevant for remote triggering because long-period waves undergo less intrinsic attenuation than do short-period waves. Directivity effects may also prove to be important[29]. Although these and other issues are speculative at this stage, the 2012 east Indian Ocean event has already shown that the phenomenon of remote earthquake triggering is not restricted to small earthquakes or tremor but extends to potentially damaging ($M \approx 7$) earthquakes. This carries implications for the effect of a large earthquake on the global seismic hazard, as the 2012 event caused strong on-land shaking (with a modified Mercalli intensity greater than or equal to seven) in Indonesia, Japan and Mexico.

METHODS SUMMARY

Determination of earthquake rates. We consider global rates, R, of shallow (depth, $\leq$100 km) $M \geq 5.5$ earthquakes during the 10 d preceding and following a given mainshock. These rates exclude pre-mainshock and post-mainshock events within a spherical cap of radius 1,500 km centred on the mainshock centroid. A running average in time is used to count events. For a target time t_{pre} preceding the event, the events are summed in bins spanning a time interval ($t_{\text{pre}} - 1$ d, min{$t_{\text{pre}} + 1$ d, 0}), where time is relative to the mainshock origin time. For a target time t_{post} following the event, the events are summed in bins spanning a time interval (max{$t_{\text{post}} - 1$ d, 0}, $t_{\text{post}} + 1$ d). This procedure is applied to the 11 April 2012 mainshock or an aggregate of $M \geq 8.5$ mainshocks in which the rates are divided by the number of mainshocks used. With the above prescription, the 2-d-average post-mainshock R value is the running-average R value evaluated at $t = 1$ d; this is the post-mainshock rate, R_{post}, used to define the seismicity rate change in equation (1).

Measure of dynamic strain. Long-period synthetic seismograms of length 2 h 33 min are calculated at 10-km depth at each of 5,150 points distributed globally for a point-source representation of the mainshock[30] convolved with a source time function consisting of a cosine ramp function of duration 100 s. They are consistent with observed seismograms (Supplementary Fig. 9). At each point, seismograms for the six components of the strain tensor are converted into time series of ε^{II}, the square root of the second invariant of the deviatoric strain tensor[31]. We define τ_{strain} to be the period during which ε^{II} is greater than a threshold value of 0.1 microstrain, using the time series of ε^{II}.

Full Methods and any associated references are available in the online version of the paper.

Received 20 May; accepted 16 August 2012.
Published online 26 September 2012.

1. Hill, D. P. *et al.* Seismicity in the western United States remotely triggered by the M7.4 Landers, California, earthquake of June 28, 1992. *Science* **260,** 1617–1623 (1993).
2. Kilb, D., Gomberg, J. & Bodin, P. Triggering of earthquake aftershocks by dynamic stresses. *Nature* **408,** 570–574 (2000).
3. Gomberg, J., Reasenberg, P., Bodin, P. & Harris, R. Earthquake triggering by transient seismic waves following the Landers and Hector Mine, California earthquakes. *Nature* **411,** 462–466 (2001).
4. Prejean, K. *et al.* Remotely triggered seismicity on the United States west coast following the M_w 7.9 Denali Fault earthquake. *Bull. Seismol. Soc. Am.* **94,** S348–S359 (2004).
5. Gomberg, J., Bodin, P., Larson, K. & Dragert, H. Earthquake nucleation by transient deformations caused by the $M = 7.9$ Denali, Alaska, earthquake. *Nature* **427,** 621–624 (2004).
6. Husen, S., Taylor, R., Smith, R. B. & Healser, H. Changes in geyser eruption behavior and remotely triggered seismicity in Yellowstone National Park produced by the 2002 M7.9 Denali Fault earthquake, Alaska. *Geology* **32,** 537–540 (2004).
7. Pankow, K. L., Arabasz, W. J., Pechman, J. C. & Nava, S. J. Triggered seismicity in Utah from the November 3, 2002, Denali Fault earthquake. *Bull. Seismol. Soc. Am.* **94,** S332–S347 (2004).
8. Brodsky, E. E. & Prejean, S. G. New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *J. Geophys. Res.* **110,** B04302 (2005).
9. Velasco, A. A., Hernandez, S., Parsons, T. & Pankow, K. Global ubiquity of dynamic earthquake triggering. *Nature Geosci.* **1,** 375–379 (2008).
10. Gonzalez-Huizar, H., Velasco, A. A., Peng, Z. & Castro, R. R. Remote triggered seismicity caused by the 2011 M9.0 Tohoku-Oki, Japan earthquake. *Geophys. Res. Lett.* **39,** L10302 (2012).
11. Parsons, T. & Velasco, A. A. Absence of remotely triggered large earthquakes beyond the mainshock region. *Nature Geosci.* **4,** 312–316 (2011).
12. Lin, C.-H. Remote triggering of the Mw 6.9 Hokkaido Earthquake as a Result of the Mw 6.6 Indonesian Earthquake on September 11, 2008. *Terr. Atmos. Ocean Sci.* **23,** 283–290 (2012).
13. McGuire, J. & Beroza, G. A rogue earthquake off Sumatra. *Science* **336,** 1118–1119 (2012).
14. Kiser, E. Preliminary Rupture Modelling of the April 11, 2012 Sumatran Earthquakes. http://www.seismology.harvard.edu/research_sumatra2012.html (2012).
15. Wang, D., Mori, J. & Ohmi, S. Rupture Process of the April 11, 2012 Sumatra (Mw 8.6) Earthquake Imaged with Back-Projection of Hi-net Data. http://www.eqh.dpri.kyoto-u.ac.jp/src/etc/sumatra.htm (2012).
16. Meng, L. Ampuero, J.-P., Duputel, Z., Luo, Y. & Tsai, V. C. Earthquake in a maze: compressional rupture branching during the 2012 M_w 8.6 Sumatra earthquake. *Science* **337,** 724–726 (2012).
17. Incorporated Research Institutions for Seismology. Back projections for MW 8.7 off W coast of Northern Sumatra. http://www.iris.edu/spud/backprojection/118733 (2012).
18. Yue, H., Lay, T. & Koper, K. D. *En échelon* and orthogonal fault ruptures of the 11 April 2012 great interplate earthquake. *Nature* http://dx.doi.org/10.1038/nature11492 (this issue).
19. Delescluse, M. *et al.* Intra-oceanic seismicity off Sumatra boosted by the Banda-Aceh megathrust. *Nature* http://dx.doi.org/10.1038/nature11520 (this issue).
20. Kanamori, H. The energy release in great earthquakes. *J. Geophys. Res.* **82,** 2981–2987 (1977).
21. Gomberg, J. & Johnson, P. Dynamic triggering of earthquakes. *Nature* **437,** 830 (2005).
22. Hill, D. P. & Prejean, S. in *Earthquake Seismology* (ed. Schubert, G.) 258–288 (Treatise on Geophysics 4, Elsevier, 2007).
23. Dahlen, F. A. & Tromp, J. *Theoretical Global Seismology* Ch. 11 (Princeton Univ. Press, 1998).
24. Gonzalez-Huizar, H. & Velasco, A. A. Dynamic triggering: stress modeling and a case study. *J. Geophys. Res.* **116,** B02304 (2011).
25. Engdahl, E. E. & Villaseñor, A. in *International Handbook of Earthquake and Engineering Seismology* (ed. Lee, W. H. K.) 665–690 (Academic, 2002).
26. Shelly, D. R., Peng, Z., Hill, D. P. & Aiken, C. Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes. *Nature Geosci.* **4,** 384–388 (2011).
27. Peng, Z. & Gomberg, J. An integrated perspective of the continuum between earthquakes and slow-slip phenomena. *Nature Geosci.* **3,** 599–607 (2010).
28. Peng, Z., Vidale, J. E., Wech, A. G., Nadeau, R. M. & Creager, K. C. Remote triggering of tremor along the San Andreas fault in central California. *J. Geophys. Res.* **114,** B00A06 (2009).
29. Velasco, A. A., Ammon, C. J., Farrell, J. & Pankow, K. Rupture directivity of the November 3, 2002 Denali Fault earthquake determined from surface waves. *Bull. Seismol. Soc. Am.* **94,** S293–S299 (2004).
30. Koss, H. & Nettles, M. Global CMT Project Moment Tensor Solution: April 11, 2012, Off W Coast of Northern Sumatra, MW = 8.6. http://earthquake.usgs.gov/earthquakes/eqinthenews/2012/usc000905e/neic_c000905e_gcmt.php (2012).
31. Jaeger, J. C. & Cook, N. G. W. *Fundamentals of Rock Mechanics* Vol. 33 (Chapman and Hall, 1984).
32. Chlieh, M. *et al.* Coseismic slip and afterslip of the Great (Mw 9.15) Sumatra-Andaman earthquake of 2004. *Bull. Seismol. Soc. Am.* **97** (1a), S152–S173 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements Epicentres and magnitudes of seismic events were obtained from the NEIC catalogue. Seismic waveform data presented in Supplementary Information were obtained from the Incorporated Research Institutions for Seismology (IRIS) Data Management Center. We thank T. Hanks, R. Harris, A. Michael, T. Parsons and P. Stark for their comments on a preliminary draft. V.S. works under contract at the US Geological Survey.

Author Contributions F.F.P. initiated the study and performed all seismic-wave analysis. F.F.P., R.S.S. and V.S. contributed equally to earthquake catalogue analysis. All authors discussed the results and helped write the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.F.P. (fpollitz@usgs.gov).

doi:10.1038/nature11420

Closing yield gaps through nutrient and water management

Nathaniel D. Mueller[1], James S. Gerber[1], Matt Johnston[1], Deepak K. Ray[1], Navin Ramankutty[2] & Jonathan A. Foley[1]

In the coming decades, a crucial challenge for humanity will be meeting future food demands without undermining further the integrity of the Earth's environmental systems[1–6]. Agricultural systems are already major forces of global environmental degradation[4,7], but population growth and increasing consumption of calorie- and meat-intensive diets are expected to roughly double human food demand by 2050 (ref. 3). Responding to these pressures, there is increasing focus on 'sustainable intensification' as a means to increase yields on underperforming landscapes while simultaneously decreasing the environmental impacts of agricultural systems[2–4,8–11]. However, it is unclear what such efforts might entail for the future of global agricultural landscapes. Here we present a global-scale assessment of intensification prospects from closing 'yield gaps' (differences between observed yields and those attainable in a given region), the spatial patterns of agricultural management practices and yield limitation, and the management changes that may be necessary to achieve increased yields. We find that global yield variability is heavily controlled by fertilizer use, irrigation and climate. Large production increases (45% to 70% for most crops) are possible from closing yield gaps to 100% of attainable yields, and the changes to management practices that are needed to close yield gaps vary considerably by region and current intensity. Furthermore, we find that there are large opportunities to reduce the environmental impact of agriculture by eliminating nutrient overuse, while still allowing an approximately 30% increase in production of major cereals (maize, wheat and rice). Meeting the food security and sustainability challenges of the coming decades is possible, but will require considerable changes in nutrient and water management.

Opportunities for agricultural intensification were analysed for seventeen major crops (which covered approximately 76% of global harvested cropland area between 1997 and 2003 (Food and Agriculture Organization of the United Nations)). Yield gaps (Fig. 1) were estimated by comparing landscape-level observed yields[12] to 'attainable yields', determined by identifying high-yielding areas within zones of similar climate. As empirical estimates, attainable yields are more conservative than absolute biophysical 'potential yields'[13], but they are probably achievable using current technology and management techniques.

Considerable yield-improvement opportunities exist relative to current attainable yield ceilings, with opportunities differing dramatically by crop and geography (regional and country-specific data for all seventeen crops are summarized in the Supplementary Information). Globally, we find that closing yield gaps to 100% of attainable yields could increase worldwide crop production by 45% to 70% for most major crops (with 64%, 71% and 47% increases for maize, wheat and rice, respectively). Eastern Europe and Sub-Saharan Africa show considerable 'low-hanging' intensification opportunities for major cereals (Fig. 2); these areas could have large production gains if yields were increased to only 50% of attainable yields. East and South Asia also have substantial intensification opportunities owing to their vast agricultural lands and the geographic variability in their yields and yield gaps.

Assessing opportunities for more sustainable intensification requires an understanding of the factors driving yield variation across the world. Fundamentally, yield gaps are caused by deficiencies in the biophysical crop growth environment that are not addressed by agricultural management practices. Here we explicitly examined key biophysical drivers of crop yield by using global, crop-specific irrigation data[14] and by developing a new global, crop-specific data set of nitrogen (N), phosphate (P_2O_5) and potash (K_2O) fertilizer application rates. We find extensive geographic variation in these management practices, with high fertilizer application rates concentrated in high-income and some rapidly developing countries (Fig. 3a and Supplementary Fig. 1). Likewise, irrigated areas[14] are heavily concentrated in South Asia, East Asia and parts of the United States (Fig. 3b).

Using input–yield crop models, we found that the spatial patterns of climate, fertilizer application and irrigated area explain 60% to 80% of global-yield variability for most major crops (Supplementary Information and Supplementary Table 1). Yields of some crops (for example, sorghum, millet and groundnut) were primarily controlled by climate, whereas others (for example, barley, sugar beet and oil palm) showed strong management responses. Surprisingly, model residuals showed little sensitivity to soil and slope parameters (Supplementary Information and Supplementary Fig. 2), suggesting that such relationships are obscured on the landscape scale with existing data sets.

The factors that primarily limit increasing crop yields to within 75% of their attainable yields (Fig. 4, Supplementary Fig. 3) vary by crop and region. For example, Eastern Europe and West Africa stand out as hotspots of nutrient limitation for maize, whereas Eastern Europe seems to experience nutrient limitation for wheat. Co-limitation of nutrients and water is observed across East Africa and Western India for maize, portions of the US Great Plains and the

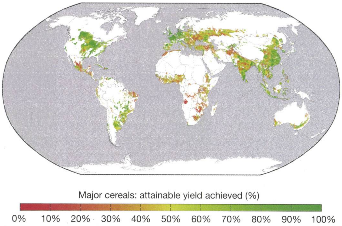


Figure 1 | Average yield gaps for maize, wheat and rice. These were measured as a percentage of the attainable yield achieved circa the year 2000. Yield gap in each grid cell is calculated as an area-weighted average across the crops and is displayed on the top 98% of growing area.

---

[1]Institute on the Environment (IonE), University of Minnesota, St. Paul, Minnesota 55108, USA. [2]Department of Geography and Global Environmental and Climate Change Center, McGill University, Montreal, Quebec H3A 2K6, Canada.

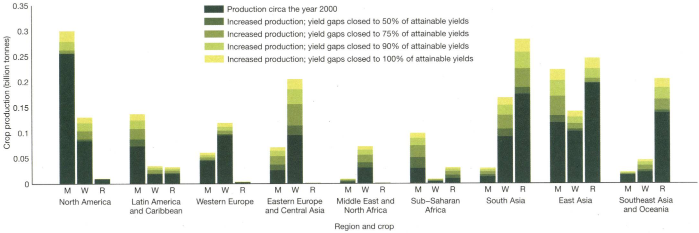


Figure 2 | Global production increases for maize, wheat and rice from closing yield gaps to 50%, 75%, 90% and 100% of attainable yields. The greatest opportunities for increases in absolute production (from closing yield gaps to 100% of estimated attainable yields) are wheat (W) in Eastern Europe and Central Asia, rice (R) in South Asia and maize (M) in East Asia. Absolute production increases for individual crops in Sub-Saharan Africa are smaller owing to lower attainable yields and diverse cropping systems (that is, less area devoted to any one crop). The region could still achieve large production increases in cassava, maize and sugarcane.

Mediterranean Basin for wheat, and in Southeast Asia for rice. We note that the management practices that limit yield increases depend on the degree of yield-gap closure desired (Supplementary Fig. 4). For example, closing maize yield gaps to 50% of attainable yields (approximately 2.5 tonnes per hectare) in Sub-Saharan Africa primarily requires addressing nutrient deficiencies (Supplementary Fig. 4a), but closing yield gaps to 75% of attainable yields (approximately 3.6 tonnes per hectare) requires increases in both irrigated area and nutrient application over most of the region (Fig. 4a).

We examined potential changes in irrigated area and nutrient application that are needed to close yield gaps of maize, wheat and rice to within 75% attainable yields (a 29% global production increase) using our input–yield models. On the landscape scale, yield gaps in co-limited regions can be closed through a range of irrigated-area and nutrient-intensity combinations (see Supplementary Fig. 5). For example, 73% of these underachieving areas could close yield gaps by solely focusing on nutrient inputs (with 18%, 16% and 35% increases in N, P_2O_5 and K_2O application relative to baseline global consumption, respectively), whereas only 16% of underachieving areas could close yield gaps by solely increasing irrigation. Jointly increasing irrigated area and nutrient application could close yield gaps on all underachieving areas (with 30%, 27% and 54% increases in N, P_2O_5 and K_2O application, respectively, and a 25% increase in irrigated hectares; Fig. 5a, b).

To minimize the environmental impacts of intensification, increased irrigation and nutrient application to close crop yield gaps should be complemented by efforts to decrease overuse of crop inputs wherever possible[15–18]; combined, these efforts could increase total food production while decreasing the overall global use of water and nutrients. For example, we estimate that by addressing imbalances and inefficiencies, nitrogen- and phosphate-fertilizer application on maize, wheat and rice could decrease globally by 11 million tonnes of nitrogen (28%) and 5 million tonnes of phosphate (38%) without impacting current yields (Supplementary Fig. 6). Nutrient overuse on these crops is particularly dramatic in China, confirming field-scale results[15]. To close yield gaps to 75% of attainable yields while also eliminating input overuse (under joint nutrient and irrigation intervention), we project that smaller net changes in nutrient inputs would be required: 9%, −2% and 34% changes in N, P_2O_5 and K_2O application (Fig. 5c and Supplementary Fig. 7). Notably, it would be possible to close global yield gaps on major cereals to within 75% of attainable yields with fairly minimal changes to total worldwide nitrogen and phosphate use by coupling targeted intensification with efforts to reduce nutrient imbalances and inefficiencies. Geographically optimizing input intensity and increasing field-scale efficiencies (beyond the average efficiencies implicit in our input–yield models) could improve production further relative to inputs.

Closing yield gaps may not always be desirable or practical in the short term, given marginal returns for additional inputs, regional

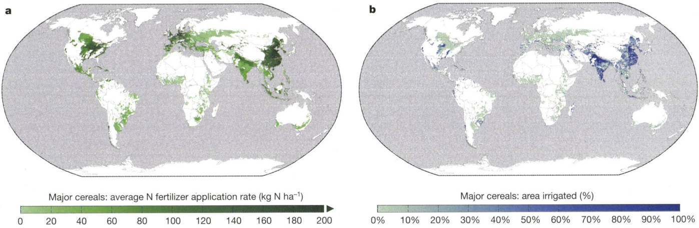

a
b
Major cereals: average N fertilizer application rate (kg N ha^{-1})
0
20
40
60
80
100
120
140
160
180
200
Major cereals: area irrigated (%)
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%

Figure 3 | Management intensity of nitrogen fertilizer and irrigated area[14] varies widely across the world's croplands. a, b, Fertilizer (**a**) and irrigation (**b**) values are area-weighted averages across major cereals.

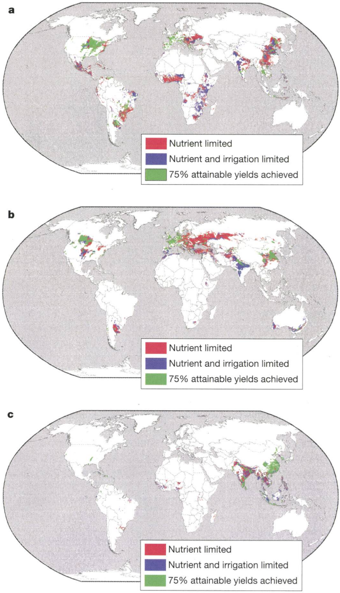

a
🟥 Nutrient limited
🟪 Nutrient and irrigation limited
🟩 75% attainable yields achieved
b
🟥 Nutrient limited
🟪 Nutrient and irrigation limited
🟩 75% attainable yields achieved
c
🟥 Nutrient limited
🟪 Nutrient and irrigation limited
🟩 75% attainable yields achieved

Figure 4 | Management factors limiting yield-gap closure to 75% of attainable yields for maize, wheat and rice. **a**, **b**, **c** Yield-limiting management factors for maize (**a**), wheat (**b**) and rice (**c**) were calculated using the suite of input–yield models, comparing current input intensity against estimated required levels to close yield gaps.

land-management policies, limits on sustainable water resources and socio-economic constraints (for example, access to capital, infrastructure, institutions and political stability). However, use of precision agriculture techniques, conservation tillage, high-yielding hybrids, increased plant populations and multifunctional landscape management can help to mitigate negative environmental impacts of intensive agriculture[19–21]. Additionally, use of organic fertilizers (omitted in this analysis owing to data limitations) are essential for improving soil carbon, enhancing soil biota and increasing water-holding capacity[22]. Social triggers of intensification will differ across regions; for example, because of development interventions by governments or NGOs, market-driven incentives for farmer investment, and land scarcity in regions not fully connected to global markets[23].

Changes to agricultural management to close yield gaps should be considered in the context of climate change, which is expected to substantially impact yields[24,25] and induce management adaptations[26]. Specifically, a major concern is how changes in water availability may conflict with projected irrigation requirements for closing yield gaps.

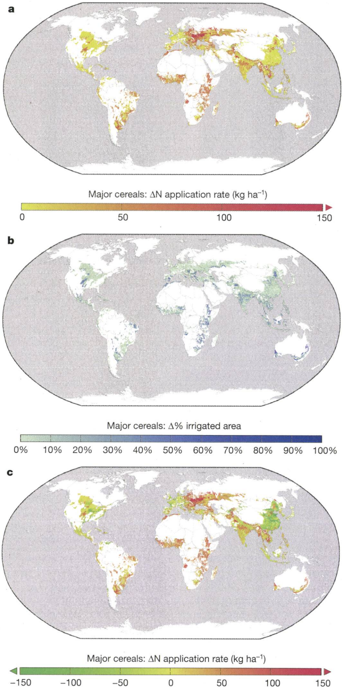

a
Major cereals: ΔN application rate (kg ha^{-1})
0
50
100
150
b
Major cereals: Δ% irrigated area
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
c
Major cereals: ΔN application rate (kg ha^{-1})
−150
−100
−50
0
50
100
150

Figure 5 | Closing yield gaps through changes in agricultural management. **a**, **b**, Projected increases in nitrogen application rates (**a**) and irrigated areas (**b**) necessary to close maize, wheat and rice yield gaps to 75% of attainable yields. **c**, Projected net changes in nitrogen application rates when closing yield gaps and eliminating input imbalances and inefficiencies.

The fertilizer data set, yield gap estimates and yield models presented here could be used widely to assess intensification opportunities and the environmental impacts of changing agricultural systems. However, these data and analyses are not without limitations (full discussion in Supplementary Information). Most importantly, the analyses rely on agricultural management, yield and climate data from a variety of different sources and on different scales. Overall, these results are most useful across regional and global scales, leaving fine-scale and temporal details obscured (for example, intra- and inter-annual variation in climate and yield creates particular uncertainty about irrigation requirements). Moreover, although our models confirm the importance of climate, fertilizers and irrigation in determining contemporary patterns of global cropland productivity, we do not discount the importance of additional biophysical characteristics (including soil characteristics, see Supplementary Information) and management practices (including crop rotation patterns, organic nutrient inputs, micronutrients, improved seed quality, conservation tillage and pest management). Incorporating these factors into the analytical framework could improve the accuracy and utility of the analyses. Additional research on cropland intensification must also assess the opportunities and environmental tradeoffs for increasing cropping intensity and decreasing pre- and post-harvest crop losses.

The future of agriculture faces two great challenges: substantial increases in food demand must be met while decreasing agriculture's global environmental footprint. Closing yield gaps and increasing resource efficiency are necessary strategies towards meeting this challenge, but they must be combined with efforts to halt agricultural expansion, reduce food waste and promote sensible diets, and produce advanced crop varieties[1,4]. This analysis emphasizes the crucial role of nutrient and water management in pathways towards sustainable intensification, and provides a starting point for a more comprehensive discussion of intensification opportunities and challenges. Context-dependent policies and agricultural development programs must address drivers of yield limitation while encouraging management practices that improve tradeoffs between production and environmental impacts.

METHODS SUMMARY

Yield gaps were quantified by comparing existing yields to climate-specific attainable yields. Our approach refines previous estimates[27,28] by excluding climate outliers and using crop-specific, equal-area climate zones.

Fertilizer application rate and consumption data were compiled for nations and subnational units across the globe (Supplementary Table 2). Application rates for crop–country combinations missing data were estimated as described in the Supplementary Information. Crop- and crop-group-specific application rates were then distributed across detailed maps of crop[12] and pasture[29] areas, and rates were harmonized with subnational and national nutrient consumption data.

Fertilizer and irrigation data were used to parameterize nutrient response curves and rainfed maximum yields, using nonlinear regression analyses within each climate zone. Using these relationships, we estimated changes in inputs necessary to close yield gaps, as well as decreases in inputs possible from addressing inefficiencies and imbalances.

Received 2 April; accepted 13 July 2012.

Published online 29 August; corrected online 10 October 2012.

1. Godfray, H. C. J. *et al.* Food security: the challenge of feeding 9 billion people. *Science* **327,** 812–818 (2010).
2. Royal Society *Reaping the benefits.* 1–86 (The Royal Society, 2009).
3. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl Acad. Sci. USA* **108,** 20260–20264 (2011).
4. Foley, J. A. *et al.* Solutions for a cultivated planet. *Nature* **478,** 337–342 (2011).
5. Robertson, G. P. & Swinton, S. M. Reconciling agricultural productivity and environmental integrity: a grand challenge for agriculture. *Front. Ecol. Environ* **3,** 38–46 (2005).
6. Cassman, K. G., Dobermann, A., Walters, D. T. & Yang, H. Meeting cereal demand while protecting natural resources and improving environmental quality. *Annu. Rev. Environ. Resour.* **28,** 315–358 (2003).
7. Foley, J. A. *et al.* Global consequences of land use. *Science* **309,** 570–574 (2005).
8. Cassman, K. G. Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Proc. Natl Acad. Sci. USA* **96,** 5952–5959 (1999).
9. Matson, P. A. & Vitousek, P. M. Agricultural intensification: Will land spared from farming be land spared for nature? *Conserv. Biol.* **20,** 709–710 (2006).
10. Clough, Y. *et al.* Combining high biodiversity with high yields in tropical agroforests. *Proc. Natl Acad. Sci. USA* 1–6 (2011).
11. Burney, J. A., Davis, S. J. & Lobell, D. B. Greenhouse gas mitigation by agricultural intensification. *Proc. Natl Acad. Sci. USA* **107,** 12052–12057 (2010).
12. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* **22,** GB1022 (2008).
13. Lobell, D. B., Cassman, K. G. & Field, C. B. Crop yield gaps: their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.* **34,** 179–204 (2009).
14. Portmann, F. T., Siebert, S. & Doell, P. MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Glob. Biogeochem. Cycles* **24,** GB1011 (2010).
15. Ju, X.-T. *et al.* Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proc. Natl Acad. Sci. USA* **106,** 3041–3046 (2009).
16. Liu, J. *et al.* A high-resolution assessment on global nitrogen flows in cropland. *Proc. Natl Acad. Sci. USA* **107,** 8035–8040 (2010).
17. MacDonald, G. K., Bennett, E. M., Potter, P. A. & Ramankutty, N. Agronomic phosphorus imbalances across the world's croplands. **108,** 3086–3091 (2011).
18. Vitousek, P. M. *et al.* Agriculture. Nutrient imbalances in agricultural development. *Science* **324,** 1519–1520 (2009).
19. Oenema, O. & Pietrzak, S. Nutrient management in food production: Achieving agronomic and environmental targets. *Ambio* **31,** 159–168 (2002).
20. Cassman, K. G., Dobermann, A. & Walters, D. T. Agroecosystems, nitrogen-use efficiency, and nitrogen management. *Ambio* **31,** 132–140 (2002).
21. Jordan, N. *et al.* Sustainable development of the agricultural bio-economy. *Science* **316,** 1570–1571 (2007).
22. Sánchez, P. A. Tripling crop yields in tropical Africa. *Nature Geosci.* **3,** 299–300 (2010).
23. Lambin, E. F. *et al.* The causes of land-use and land-cover change: moving beyond the myths. *Glob. Environ. Change* **11,** 261–269 (2001).
24. Parry, M., Canziani, O., Palutikof, J., &. co-authors *Technical Summary. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. and Hanson, C. E.) 23–78 (Cambridge Univ. Press, 2007).
25. Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **333,** 616–620 (2011).
26. Howden, S. M. *et al.* Adapting agriculture to climate change. *Proc. Natl Acad. Sci. USA* **104,** 19691–19696 (2007).
27. Licker, R. *et al.* Mind the gap: how do climate and agricultural management explain the "yield gap" of croplands around the world? *Glob. Ecol. Biogeogr.* **19,** 769–782 (2010).
28. Johnston, M. *et al.* Closing the gap: global potential for increasing biofuel production through agricultural intensification. *Environ. Res. Lett.* **6,** 034028 (2011).
29. Ramankutty, N., Evan, A. T., Monfreda, C. & Foley, J. A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **22,** GB1003 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Allez and K. Milligan for assistance with data collection. We are grateful to R. Licker, G. MacDonald, M. Mueller, S. Polasky, P. Potter, P. Reich, L. Schulte-Moore, D. Tilman, J. Van Wart and the Foley Laboratory for helpful conversations. We thank P. Robertson and P. Smith for helpful comments on the manuscript. Funding was provided by a National Science Foundation Graduate Research Fellowship and a University of Minnesota College of Food, Agricultural and Natural Resource Sciences Fellowship to N.D.M., a Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Grant to N.R., Gordon and Betty Moore Foundation funding to J.A.F., and support from the University of Minnesota Institute on the Environment.

Author Contributions N.D.M. led the study design, data analysis and writing. J.S.G. contributed substantially to the yield-gap data analysis and writing. D.K.R. and M.J. assisted with data analysis and writing. J.A.F. and N.R. assisted with study design and writing.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.D.M. (muell512@umn.edu).

LETTER

doi:10.1038/nature11495

Complex brain and optic lobes in an early Cambrian arthropod

Xiaoya Ma[1,2], Xianguang Hou[1], Gregory D. Edgecombe[2] & Nicholas J. Strausfeld[3]

The nervous system provides a fundamental source of data for understanding the evolutionary relationships between major arthropod groups[1,2]. Fossil arthropods rarely preserve neural tissue. As a result, inferring sensory and motor attributes of Cambrian taxa has been limited to interpreting external features, such as compound eyes[3] or sensilla decorating appendages[4], and early-diverging arthropods have scarcely been analysed in the context of nervous system evolution. Here we report exceptional preservation of the brain and optic lobes of a stem-group arthropod from 520 million years ago (Myr ago), *Fuxianhuia protensa*[5], exhibiting the most compelling neuroanatomy known from the Cambrian. The protocerebrum of *Fuxianhuia* is supplied by optic lobes evidencing traces of three nested optic centres serving forward-viewing eyes. Nerves from uniramous antennae define the deutocerebrum, and a stout pair of more caudal nerves indicates a contiguous tritocerebral component. *Fuxianhuia* shares a tripartite pre-stomodeal brain and nested optic neuropils with extant Malacostraca and Insecta[2,6], demonstrating that these characters were present in some of the earliest derived arthropods. The brain of *Fuxianhuia* impacts molecular analyses that advocate either a branchiopod-like ancestor of Hexapoda[7,8] or remipedes and possibly cephalocarids as sister groups of Hexapoda[9,10]. Resolving arguments about whether the simple brain of a branchiopod approximates an ancestral insect brain or whether it is the result of secondary simplification has until now been hindered by lack of fossil evidence. The complex brain of *Fuxianhuia* accords with cladistic analyses on the basis of neural characters, suggesting that Branchiopoda derive from a malacostracan-like ancestor but underwent evolutionary reduction and character reversal of brain centres that are common to hexapods and malacostracans[2]. The early origin of sophisticated brains provides a probable driver for versatile visual behaviours, a view that accords with compound eyes from the early Cambrian that were, in size and resolution, equal to those of modern insects and malacostracans[3].

Fuxianhuia protensa is one of more than 80 arthropod species retrieved from the Chengjiang biota (Cambrian Series 2, Stage 3) at various sites in Yunnan Province, southwest China[11]. Fossils of *Fuxianhuia* are commonly preserved dorsoventrally flattened (Fig. 1). Dorsal aspects show a 'head shield' comprising two hemi-shields together covering the anteriormost four segments, the first of which bears a pair of postocular uniramous antennae[12] (Fig. 2a). Ventrally, the first post-antennal segment provides paired curved articulated structures that have been identified either as appendages flanking the mouth[5,13] or as gut diverticuli[12,14]. A thorax comprises up to 16 segments, with fewer tergites than biramous appendages[5,12]. The thoracic segments have broad paratergal folds. An abruptly narrowed abdomen consists of 13 segments, and a terminal spine is flanked by a pair of shorter spines. The body of *F. protensa* is thus extremely simple, yet evidences obvious tagmosis (Fig. 1). Frontally, each of a pair of eyes is situated at the end of a short wide stalk that broadens medially to merge with a rectangular optic capsule (Fig. 2a, c), sometimes claimed as pre-segmental[12], that lies at the midline beneath a forward-thrusting rostrum, also referred to as an anterior sclerite[13]. Paired protrusions, situated medial to each eye and extending forward dorsally, were once interpreted as a second or ancillary pair of eyes[12]. Although in some fossils these protrusions are the same colour as the lateral eye (Fig. 3a), in others the colouration matches that of the rostrum (Fig. 3c). No corneal facets are preserved and, accordingly, we interpret these structures as lateral extensions of the rostrum and not compound eyes. Observations of the tergites in specimens preserved ventral side up show little indication of sensory bristles or setae on their surfaces, whereas thoracic endopods and the paired antennae show occasional insertion points of sensory setae on each article and annulus (Supplementary Fig. 1a). Comparable sensilla have been identified in the Burgess Shale arthropod *Waptia fieldensis*, indicating abundant sensory inputs to segmental ganglia[4].

As is typical of Chengjiang fossils, internal organs of *Fuxianhuia* show exceptional preservation, delineated by contours, as in the digestive tract, and often by coloured delineations from diagenetic mineralization that effectively stains or outlines preserved internal structures[12,15]. These can include segmentally arranged digestive glands/diverticuli, retinas and

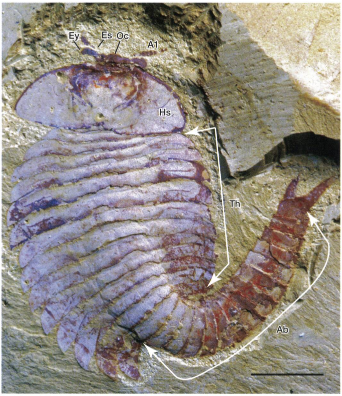


Figure 1 | *Fuxianhuia protensa* from the Chengjiang Lagerstätte. Dorsal view of complete specimen, YKLP 11321. A1, antenna; Ab, abdomen; Es, eye stalk; Ey, eye; Hs, head shield; Oc, optic capsule; Th, thorax. Scale bar, 1 cm.

---

[1]Yunnan Key Laboratory for Palaeobiology, Yunnan University, Kunming 650091, China. [2]Department of Earth Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. [3]Department of Neuroscience and Center for Insect Science, University of Arizona, Tucson, Arizona 85721, USA.

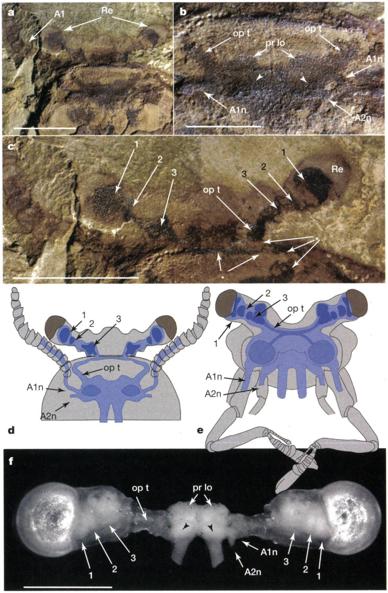


Figure 2 | Brain of *Fuxianhuia protensa* and comparison with Malacostraca. a–c, YKLP 15006. **a,** Head region, showing paired eyes (Re) and an antenna (A1). **b,** Detail of brain. Arrowheads indicate paired darker regions at level of entry of antennal nerves. **c,** Optic capsule, dark areas (arrowed 1–3) corresponding to nested optic neuropils. Arrows leading from asterisks indicate contiguous optic tract remnants, including heterolateral component within main body. **d,** Reconstructed brain of *Fuxianhuia* showing organization of optic lobes, optic tracts and peripheral nerve tracts A1n and A2n. **e,** Reconstructed brain of the land hermit crab *Coenobita clypeatus* (ref. 25) showing homologies in Malacostraca. Central projections of tracts from optic stalks, including heterolateral connections between the stalks, are shown as an intermediate shade of blue; in *Fuxianhuia* heterolateral connections are provided by a separate tract extending between the optic tracts. **f,** Top-down view of brain of glass shrimp, *Palaeomonetes pugio*. Transparent cuticle shows three nested neuropils (arrowed 1–3) in eye stalks. Protocerebrum denoted by its paired lobes (pr lo). Black arrowheads indicate olfactory lobes. Nerve roots of A1n and A2n to right. A1n, antennal nerve; A2n, second pair of nerve roots; op t, optic tract; pr lo, protocerebral lobes. Scale bars, 1 mm.

lenses, and, in rare instances, traces of muscle and nervous system. This last structure is arguably the densest internal tissue in any extant arthropod, comprising vast numbers of lipid profiles, cytoskeletal proteins and mitochondria. Profiles of internal structures, matching the profiles of contemporary brains and ganglia, have previously been identified in the Burgess Shale[4] and Chengjiang[12,16], belying the common supposition that neural tissue is too prone to decay to withstand fossilization.

Observations of some 50 specimens, many newly collected and stored at the Yunnan Key Laboratory for Palaeobiology (YKLP), demonstrate well-preserved eyes and eye stalks in various orientations (Figs 2c and 3), suggesting that the eye stalk assemblage possessed a considerable degree of rotational freedom and thus allowed active vision (Supplementary Fig. 3). For example, the eye stalk in Fig. 3a is tilted forwards, such that the eye appears as viewed from directly above, demonstrating a much larger radius of curvature frontally than laterally. Such differences of curvature are commonplace in extant insects and crustaceans, the flatter surfaces allowing small divergence angles between adjacent ommatidia and thus greater resolution of the frontal visual field than the lateral one[17]. In the closely allied *Chengjiangocaris longiformis*, the medial protrusion and the lateral eye base are associated with a Y-shaped structure on the ventral side that possibly afforded mechanical support and articulation (Fig. 3b). Eyes viewed from their frontal aspect—that is, preserved tilted upwards as in Fig. 3c—show the presence of a compound eye comprising facets. A thin opalescent layer (Fig. 3d, e) overlies two deeper stratifications interpreted as corresponding to dioptric arrangements: corneal lenses surmounting crystalline cones. Patches of lenses identified in exceptionally preserved specimens resolve a regular arrangement of 25 µm diameter lenses on both the lateral and frontal eye margins (Fig. 3f). These layered arrangements correspond to those observed in malacostracan crustaceans (Fig. 3g).

One specimen (YKLP 15006), previously described in ref. 12, shows exceptionally preserved traces of tissue within the eye stalks, and tract-like extensions from these to a bilaterally symmetric area situated frontally, beneath the head shield (Fig. 2a), identified as the brain. Its shape (Fig. 2b) matches that of a comparable-sized modern malacostracan (Fig. 2f). The brain in Fig. 2b shows paired anterior protocerebral lobes, behind which the tracts from the eye stalks converge. Caudal to these emerge robust traces of a pair of antennal nerves, one of which (A1n) can be followed into the antenna (Fig. 2b). Immediately behind these nerves is a third pair of stout tracts. These do not emanate from the antennae, but probably belong to post-antennal appendages. The presence of these paired nerve roots (A2n) supports observations identifying the paired sickle-shaped structures flanking the mouth ventrally as appendicular[5,13]. The brain of YKLP 15006 shows definitive bilateral symmetry both in its outlines and in the three pairs of nerves supplying it. A pair of symmetrically positioned darker textured

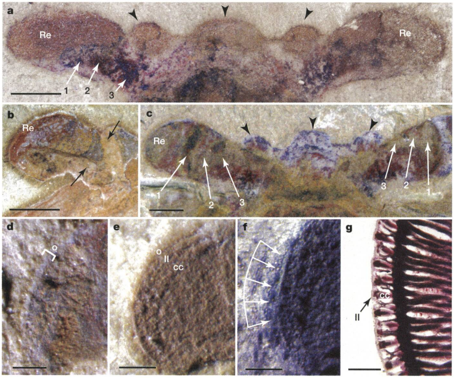


Figure 3 | Eyes of fuxianhuiids and comparison to Malacostraca. **a**, Paired eyes of *Fuxianhuia protensa*, YKLP 11322. Top-down view of left eye stalk showing retina (Re), flanking protrusion and midline rostral sclerite (arrowheads). Dark areas (arrowed 1–3) suggest three nested neuropils. **b**, Ventral view of eye of *Chengjiangocaris longiformis*, RCCBYU 10256, showing Y-shaped sclerite extending from ventral surface of first segment. **c–f**, *Fuxianhuia protensa*, YKLP 11323. **c**, Exposed interior of eye stalk, resolving three raised, nested areas. **d**, Right eye margin showing opalescent layer interpreted as the corneal surface (o). **e**, Left eye, two layers beneath outer layer correspond to the lens layer (ll) and crystalline cone (cc), together providing the dioptric layer. **f**, Left eye and its margins showing regularly spaced structures corresponding to lenses. **g**, Silver-stained lateral eye margin of a malacostracan (*Lebbeus groenlandicus*) showing argyrophilic lens layer beneath a thin corneal layer and a layer of crystalline cones. Scale bars, **a–d** 250 µm, **e–h** 50 µm.

domains is observed within the brain at the level of entry of the antennal nerves (Fig. 2b). The position of these paired regions best corresponds to the location of paired olfactory lobes observed in the deutocerebrum of numerous crown-group malacostracans[6,18].

The eye stalks of YKLP 15006 and other specimens contain alternating swollen and constricted patches of pigmentation that are interpreted as volumes of optic neuropils. Their identity as neural tissue is substantiated by the continuity of the optic tract with the anterior part of the brain. Three successive domains of pigmentation are indicative of three nested neuropils (numbered as in Figs 2c and 3c and Supplementary Figs 1 and 2). Depending on the orientation of the head, these domains can appear to be positioned towards the rear margin of the eye stalk (Fig. 2c) or further forward when the eyes and head capsule are rotated upwards around the horizontal axis (Fig. 3c and Supplementary Fig. 2a, b). Neural tissue is shown as dark material (Fig. 2a–c) or the profiles of optic regions and optic tracts may be the only areas lacking diagenetic deposits, as shown symmetrically in both eye stalks in Fig. 3c (Supplementary Fig. 2a, b). In YKLP 15006, the putative optic neuropils are contiguous with darkly coloured optic nerves that extend caudally to eventually merge into the brain area at the level of its anterior protocerebral lobes (Fig. 2c and Supplementary Fig. 1). Scanning electron microscopy and energy-dispersive X-ray spectroscopy show that both the brain and the optic lobes are preserved with higher iron concentration than other surrounding structures, including the cuticle (Supplementary Fig. 4).

The presence of three successive pairs of nerve roots, corresponding to the optic nerves, paired uniramous antennae and evidently to paired post-antennal appendages, demonstrates that the brain of *Fuxianhuia* comprises three neuromeres: a protocerebrum, deutocerebrum and tritocerebrum. It is not possible to ascertain whether the entire tritocerebral ganglion lies in front of the stomodeum or whether, as in many extant pancrustaceans, the tritocerebrum is penetrated by the stomodeum, with one part of it pre-stomodeal and the other part post-stomodeal. In other mandibulate arthropods, such as centipedes, the tritocerebrum flanks the gut, but its neuropil is contiguous with, not separate from, the deutocerebrum[19]. In no taxon is the tritocerebrum wholly pre-oral: in cases where the entire tritocerebrum appears to lie forward of the gut, there is still one bundle of post-stomodeal axons that connects both sides of the tritocerebrum[20].

The brain of *Fuxianhuia* is thus tripartite, as are brains of Malacostraca[18], Chilopoda[19] and Insecta[20], but not those of Branchiopoda, in which the tritocerebral ganglion is separate from the deutocerebrum and situated post-orally[21]. Likewise, the indication of three nested optic neuropils in *Fuxianhuia* (Fig. 2) contrasts with Branchiopoda, which have only two such neuropils linked by uncrossed axons[22], rather than three nested neuropils linked by chiasmata, as in malacostracans, including Phyllocarida, and insects[23,24] (Fig. 2f). The reconstructed brain of *Fuxianhuia* (Fig. 2d) thus corresponds in many respects to that of malacostracans[25] (Fig. 2e).

The crucial differences between the branchiopod and malacostracan central nervous system are evidenced in other important ways. Branchiopoda lack olfactory lobes and hemiellipsoid bodies, both of which typify not only Malacostraca but also Remipedia[26] and Copepoda[27]. Even Cephalocarida possess glomerular antennal lobes that are connected to paired mushroom bodies, the homologues of hemiellipsoid bodies[25,28]. The elaborate mid-line neuropils of Malacostraca, consisting of a protocerebral bridge and bilateral central body, contrast with the simpler central body in Branchiopoda[2,6]. The more simple arrangement of two retinotopic neuropils connected by uncrossed axons has been used, classically[29], to advocate branchiopods as basal within Pancrustacea (=Tetraconata *sensu* ref. 30), diverging before a more elaborate tripartite brain and additional optic neuropil (the lobula) evolved in Malacostraca and Insecta. If Branchiopoda indeed show plesiomorphic character states, fossil brains from deep time ought to provide supporting evidence. Here, observations of the fossilized brain of a putative stem-group arthropod, *Fuxianhuia*, instead provides evidence of a tripartite brain anterior to the stomodeum, supplied by optic nerves originating from lobes that appear to comprise three nested neuropils, indicating that, if homologous, these characters may pre-date the most recent common ancestor of extant Pancrustacea.

The absence of specialized head appendages posterior to the paired curved post-antennal appendages[5] and the structure of the trunk appendages (with simple exopod flaps and no differentiation of a protopodite) have argued for a position of *Fuxianhuia* in the arthropod stem-group[5,12–14]; that is, branching earlier than the most recent common ancestor of Chelicerata and Mandibulata (phylogeny as in ref. 10). The brain of *Fuxianhuia* is more readily compared to that of mandibulates than to chelicerates, which, apart from Araneae, have only one bona fide retinotopic neuropil and lack reversing chiasmata and optic lobes per se. Spider visual systems entirely separate the colour-sensitive and colour-insensitive pathways and have convergently evolved two nested retinotopic neuropils for each eye type, plus two deeper structures that are most similar to central complex neuropils and mushroom bodies[6].

A position of *Fuxianhuia* in the arthropod stem-group requires chelicerate neural characters to be simplified compared with an ancestor that possessed the shared features of *Fuxianhuia* and crown pancrustaceans, excluding Branchiopoda, and it implies that the ground pattern organization of the crustacean brain evolved very early. Alternatively, it could imply that the malacostracan-like characters of *Fuxianhuia* are convergent with those of Pancrustacea/Tetraconata. In either case, the brain and optic lobes of *Fuxianhuia* suggest that the arthropod nervous system acquired complexity by the early Cambrian. Indeed, it is expected that optic lobes would have already evolved sophisticated circuits even more deeply in the arthropod stem-group, enabling high-level visual processing of the kind presumed to be associated with large compound eyes belonging to the stem-group arthropod *Anomalocaris*[3].

METHODS SUMMARY

All studied specimens are deposited in the Yunnan Key Laboratory for Palaeobiology, Yunnan University, Kunming, China. Specimens were prepared with fine needles under high magnification using stereomicroscopes. Digital images of the fossils were captured using a Nikon SMZ1000 photomicroscope and were processed in Adobe Photoshop CS 5. Figure 2f was taken using a Zeiss Stemi SV11 photomicroscope equipped with a Jenoptic ProgRes C3 digital camera. Image enhancement of lens periodicity (Fig. 3f) was obtained by superimposing two identical images of the eye, colour substituted for contrast, and using the darkening function to shift the top image upwards by a distance equal to the width of one lens. The image of Fig. 3a had green tones removed to accentuate contrast of coloured structures against the matrix.

Full Methods and any associated references are available in the online version of the paper.

Received 18 June; accepted 9 August 2012.

1. Harzsch, S. Neurophylogeny: architecture of the nervous system and a fresh view on arthropod phylogeny. *Integr. Comp. Biol.* **46,** 162–194 (2006).
2. Strausfeld, N. J. & Andrew, D. R. A new view of insect–crustacean relationships I. Inferences from neural cladistics and comparative neuroanatomy. *Arthropod Struct. Dev.* **40,** 276–288 (2011).
3. Paterson, J. R. *et al.* Acute vision in the giant Cambrian predator *Anomalocaris* and the origin of compound eyes. *Nature* **480,** 237–240 (2011).
4. Strausfeld, N. J. Some observations on the sensory organization of the crustaceamorph *Waptia fieldensis* Walcott. *Palaeontogr. Canadiana* **31,** 157–169 (2011).
5. Chen, J.-Y., Edgecombe, G. D., Ramsköld, L. & Zhou, G.-Q. Head segmentation in early Cambrian *Fuxianhuia*: implications for arthropod evolution. *Science* **268,** 1339–1342 (1995).
6. Strausfeld, N. J. *Arthropod Brains: Evolution, Functional Elegance, and Historical Significance* (Belknap Press, 2012).
7. Glenner, H., Thomsen, P. F., Hebsgaard, M. B., Sørensen, M. V. & Willerslev, E. The origin of insects. *Science* **314,** 1883–1884 (2006).
8. Campbell, L. I. *et al.* MicroRNAs and phylogenomics resolve the relationships of Tardigrada and suggest that velvet worms are the sister group of Arthropoda. *Proc. Natl Acad. Sci. USA* **108,** 15920–15924 (2011).
9. von Reumont, B. M. *et al.* Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol. Biol. Evol.* **29,** 1031–1045 (2012).
10. Regier, J. C. *et al.* Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463,** 1079–1083 (2010).
11. Hou, X.-G. *et al. The Cambrian Fossils of Chengjiang, China: The Flowering of Early Animal Life* (Blackwell, 2004).
12. Bergström, J., Hou, X.-G., Zhang, X.-G. & Clausen, S. A new view of the Cambrian arthropod *Fuxianhuia. GFF* **130,** 189–201 (2008).
13. Budd, G. E. Head structure in upper stem-group euarthropods. *Palaeontology* **51,** 561–573 (2008).
14. Waloszek, D., Chen, J.-Y., Haas, A. & Wang, X.-Q. Early Cambrian arthropods – new insights into arthropod head and structural evolution. *Arthropod Struct. Dev.* **31,** 189–205 (2005).
15. Gaines, R. R. *et al.* Mechanism for Burgess Shale-type preservation. *Proc. Natl Acad. Sci. USA* **109,** 5180–5184 (2012).
16. Hou, X.-G., Ma, X.-Y., Zhao, J. & Bergstrom, J. The lobopodian *Paucipodia inermis* from the lower Cambrian Chengjiang fauna, Yunnan, China. *Lethaia* **37,** 234–244 (2004).
17. Land, M. F. in *Facets of Vision* 90–111 (Springer, 1989).
18. Sandeman, D. C. & Scholtz, G. in *The Nervous Systems of Invertebrates: An Evolutionary and Comparative Approach* 329–347 (Birkhäuser, 1995).
19. Sombke, A. *et al.* Comparative analysis of deutocerebral neuropils in Chilopoda (Myriapoda): implications for the evolution of the arthropod olfactory system and support for the Mandibulata concept. *BMC Neurosci.* **13,** 1–17 (2012).
20. Boyan, G. S., Williams, J. L. D. & Hirth, F. in *Evolution of Nervous Systems I* 249–360 (Academic Press, 2006).
21. Kirsch, R. & Richter, S. The nervous system of *Leptodora kindtii* (Branchiopoda, Cladocera) surveyed with confocal scanning microscopy (CLSM), including general remarks on the branchiopod neuromorphological ground pattern. *Arthropod Struct. Dev.* **36,** 143–156 (2007).
22. Elofsson, R. & Dahl, E. The optic neuropils and chiasmata of Crustacea. *Z. Zellforsch. Mikrosk. Anat.* **107,** 343–360 (1970).
23. Strausfeld, N. J. The evolution of crustacean and insect optic lobes and the origin of chiasmata. *Arthropod Struct. Dev.* **34,** 235–256 (2005).
24. Sinakevitch, I., Douglass, J. K., Scholtz, G., Loesel, R. & Strausfeld, N. J. Conserved and convergent organization in the optic lobes of insects and isopods, with reference to other crustacean taxa. *J. Comp. Neurol.* **467,** 150–172 (2003).
25. Wolff, G., Harzsch, S., Hansson, B. S., Brown, S. & Strausfeld, N. Neuronal organization of the hemiellipsoid body of the land hermit crab *Coenobita clypeatus*: correspondence with the mushroom body ground pattern. *J. Comp. Neurol.* **520,** 2824–2846 (2012).
26. Fanenbruck, M. & Harzsch, S. A brain atlas of *Godzilliognomus frondosus* Yager, 1989 (Remipedia, Godzilliidae) and comparison with the brain of *Speleonectes tulumensis* Yager, 1987 (Remipedia, Speleonectidae): implications for arthropod relationships. *Arthropod Struct. Dev.* **34,** 343–378 (2005).
27. Andrew, D. R., Brown, S. M. & Strausfeld, N. J. The minute brain of the copepod *Tigriopus californicus* supports a complex ancestral ground pattern of the tetraconate cerebral nervous systems. *J. Comp. Neurol.* **520,** 3446–3470 (2012).
28. Stegner, M. E. & Richter, S. Morphology of the brain in *Hutchinsoniella macracantha* (Cephalocarida, Crustacea). *Arthropod Struct. Dev.* **40,** 221–243 (2011).
29. Hanström, B. Eine genetische Studie über die Augen und Sehzentren von Turbellarien, Anneliden und Arthropoden (Trilobiten, Xiphosuren, Eurypteriden, Arachnoiden, Myriapoden, Crustaceen und Insekten). *Kungl. Svensk. Vetenskapsakad. Handl.* **4,** 1–176 (1926).
30. Dohle, W. Are the insects terrestrial crustaceans? A discussion of some new facts and arguments and the proposal of the proper name 'Tetraconata' for the monophyletic unit Crustacea + Hexapoda. *Ann. Soc. Entomol. Fr.* **37,** 85–103 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements This account is dedicated to J. Bergström, who was first in suggesting a brain in YKLP 15006. The present work was supported by the National Natural Science Foundation of China (40730211), a Leverhulme Trust Research Project Grant (F/00 696/T), and the Center for Insect Science, University of Arizona. We thank A. Ball and J. Spratt for their assistance with scanning electron microscopy imagery and energy-dispersive X-ray spectroscopy, respectively.

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to X.M. (x.ma@nhm.ac.uk).

doi:10.1038/nature11466

Dopaminergic neurons inhibit striatal output through non-canonical release of GABA

Nicolas X. Tritsch[1], Jun B. Ding[1]† & Bernardo L. Sabatini[1]

The substantia nigra pars compacta and ventral tegmental area contain the two largest populations of dopamine-releasing neurons in the mammalian brain. These neurons extend elaborate projections in the striatum, a large subcortical structure implicated in motor planning and reward-based learning. Phasic activation of dopaminergic neurons in response to salient or reward-predicting stimuli is thought to modulate striatal output through the release of dopamine to promote and reinforce motor action[1–4]. Here we show that activation of dopamine neurons in striatal slices rapidly inhibits action potential firing in both direct- and indirect-pathway striatal projection neurons through vesicular release of the inhibitory transmitter GABA (γ-aminobutyric acid). GABA is released directly from dopaminergic axons but in a manner that is independent of the vesicular GABA transporter VGAT. Instead, GABA release requires activity of the vesicular monoamine transporter VMAT2, which is the vesicular transporter for dopamine. Furthermore, VMAT2 expression in GABAergic neurons lacking VGAT is sufficient to sustain GABA release. Thus, these findings expand the repertoire of synaptic mechanisms used by dopamine neurons to influence basal ganglia circuits, show a new substrate whose transport is dependent on VMAT2 and demonstrate that GABA can function as a bona fide co-transmitter in monoaminergic neurons.

The striatum integrates inputs from cortex, hippocampus, thalamus, amygdala and ventral tegmental area/substantia nigra pars compacta (VTA/SNc) to instruct the selection of appropriate motor actions. Inputs from midbrain dopamine (DA) neurons play an important role in this process, as evidenced by the psychomotor deficits that arise after loss of these cells in Parkinson's disease, or by the occurrence of compulsive and addictive behaviours upon potentiation of dopaminergic signalling[5–7]. Through release of DA, these neurons promote activation of direct-pathway striatal projection neurons (dSPNs), which express $G\alpha_{s/olf}$-coupled D_1 receptors, and inhibit indirect-pathway SPNs (iSPNs), which express $G\alpha_{i/o}$-coupled D_2 receptors[3,5]. However, midbrain DA neurons also express neuropeptides[8] and a subset releases glutamate[9–12], suggesting that the net effects of activity in these cells may not be limited to the actions of DA.

To investigate how DA neurons influence neuronal activity in striatum, we expressed the light-activated cation channel channelrhodopsin-2 (ChR2)[13] in SNc neurons using Cre recombinase-dependent adeno-associated viruses (AAVs). In *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$ mice, Cre expression shows high penetrance and specificity for midbrain DA neurons[14] (Supplementary Fig. 1). We validated SNc targeting using an AAV encoding Cre-dependent enhanced green fluorescent protein (EGFP). Fluorescence was restricted to Cre-containing neurons in SNc, and most EGFP-expressing cells (97.9%, $n = 1587$ cells; three mice) were immunopositive for the catecholamine biosynthetic enzyme tyrosine hydroxylase, demonstrating specific expression in DA neurons (Fig. 1a, b and Supplementary Figs 1 and 2). Moreover, EGFP^+ axons densely innervated dorsal striatum (Fig. 1c), consistent with their nigrostriatal identity.

Carbon-fibre amperometry confirmed the ability to evoke DA release from ChR2-expressing axons in slices of dorsal striatum. Brief flashes of blue light (1 ms) reliably evoked DA transients, which were sensitive to the tyrosine hydroxylase antagonist α-methyl-tyrosine and

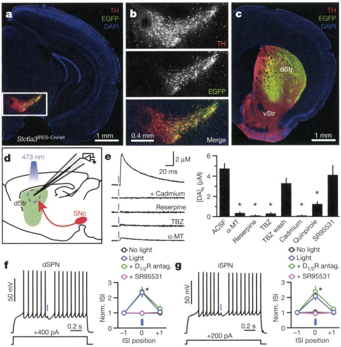


Figure 1 | DA neuron stimulation inhibits SPNs. **a**, Coronal midbrain section from a *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$ mouse transduced with a Cre-dependent EGFP AAV (green). DA neurons immunolabelled for tyrosine hydroxylase (TH, red). DAPI nuclear stain (blue). **b**, Higher magnification of boxed area in **a**. Although expression levels vary greatly, most EGFP^+ cells are tyrosine-hydroxylase-positive. **c**, EGFP^+ SNc neurons densely innervate dorsal striatum (dStr). vStr; ventral striatum. **d**, Carbon-fibre recording configuration in a sagittal brain slice. $\mathrm{ChR2}^+$ DA neurons depicted red, laser illumination area blue. **e**. Left, light stimulation (blue) of DA terminals in dorsal striatum evoked cadmium-, reserpine-, TBZ- and α-methyl-tyrosine (α-MT)-sensitive DA release measured by amperometry. Stimulation artefacts blanked for clarity. Right, mean ($n = 4$–19) peak extracellular DA concentration after optogenetic stimulation of nigrostriatal axons. DA transients recovered upon TBZ washout (wash). $^*P < 0.05$ versus ACSF (Mann–Whitney rank sum test). **f**, **g**, Left, membrane potential responses of a dorsal striatum dSPN (**f**) or iSPN (**g**) to current injection (1.3 s, bottom) and ChR2-mediated stimulation of DA axons (1 ms, blue). Right, average normalized duration of three consecutive interspike intervals (Norm. ISI) that precede, straddle or follow the light flash (positions '−1', '0' and '+1', respectively) upon ChR2 stimulation in ACSF (blue), DA receptor blockers ($D_{1/2}$R antagonists = SCH23390 + SKF83566 + sulpiride + L-741,626; green) or SR95531 ($\mathrm{GABA_A}$ receptor blocker; purple). Black, light prevented from entering the sample; $n = 14$ dSPNs, 11 iSPNs. $^*P < 0.05$ (two-way ANOVA). Data in **e**–**g** represent mean ± s.e.m.

---

[1]Department of Neurobiology, Howard Hughes Medical Institute, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA. †Present address: Department of Neurology and Neurological Sciences, Stanford Institute of Neuro-innovation and Translational Neurosciences (SINTN), Stanford University School of Medicine, 1050 Arastradero Road, Palo Alto, California 94304, USA.

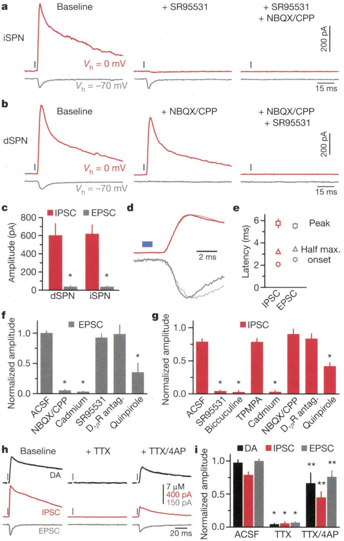


Figure 2 | DA neurons directly release GABA onto SPNs. **a**, Evoked current responses from an iSPN held at indicated potentials (V_h) to optogenetic activation of nigrostriatal axons (1 ms, blue) upon sequential bath application of SR95531 and NBQX/CPP. **b**, As in **a** for a dSPN with antagonists applied in reverse order. **c**, Mean IPSC (red) and EPSC (grey) absolute amplitudes in dSPNs ($n = 8$) and iSPNs ($n = 21$). *$P < 0.05$ versus IPSC (Mann–Whitney rank sum test). **d**, Normalized IPSCs and EPSCs from **a** (dark) and **b** (light) shown on an expanded timescale. Blue, light presentation. **e**, Average ($n = 29$ SPNs) IPSC (red) and EPSC (grey) latencies from light onset to current onset (circle), half maximal amplitude (triangle) and peak amplitude (square). IPSCs are not delayed compared with EPSCs. **f**, **g**, Mean ($n = 3$–18) EPSC (**f**) and IPSC (**g**) amplitudes under control conditions (ACSF) or in indicated antagonists normalized to baseline. *$P < 0.05$ versus ACSF (Mann–Whitney rank sum test). **h**, Amperometric DA transients (black) and current responses of a voltage-clamped iSPN ($V_h = 0$ mV, red; $V_h = -70$ mV, grey) to DA neuron stimulation under baseline conditions (left), in TTX (middle) and after co-application of TTX and 4-aminopyridine (4AP, right). **i**, Extracellular DA concentration (black), IPSC amplitude (red) and EPSC amplitude (grey) evoked by ChR2 stimulation across conditions normalized to baseline. Data from ACSF condition same as in Figs 1e and 2f, g. *$P < 0.05$ versus ACSF; ** < 0.05 versus TTX (Mann–Whitney rank sum test). Error bars, s.e.m.

the VMAT2 antagonists reserpine and tetrabenazine (TBZ) (Fig. 1d, e). Moreover, DA release required Ca^{2+} channels, as it was blocked by cadmium, and activation of D_2 receptors with quinpirole reduced DA release, consistent with the effect of DA autoreceptors in nigrostriatal axons[15].

To determine the net effect of DA neuron stimulation on striatal output, we performed whole-cell current-clamp recordings from dSPNs and iSPNs in brain slices obtained from $\mathit{Slc6a3}^{\mathrm{IRES-Cre/wt}}$;*Drd2*-EGFP mice. Whereas action potentials evoked by current steps occurred at regular intervals under baseline conditions, light presentation reliably paused firing and rapidly hyperpolarized SPNs of both populations (dSPNs: 6.8 ± 1.9 mV, $n = 10$; iSPNs: 7.0 ± 1.4 mV, $n = 7$) (Fig. 1f, g). The light-evoked pause and hyperpolarization were unaffected by a cocktail of antagonists targeting DA receptors (dSPNs: 6.6 ± 2.2 mV, $n = 4$; iSPNs: 9.5 ± 2.5 mV, $n = 4$; both $P > 0.05$ versus light only, Mann–Whitney rank sum test). Instead, they were abolished by the $GABA_A$ receptor antagonist SR95531 ($n = 4$ dSPNs, seven iSPNs) (Fig. 1f, g), which does not alter DA release (Fig. 1e). These data indicate that DA neurons exert a rapid and strong inhibitory influence on SPNs through activation of $GABA_A$ receptors.

Previous attempts at characterizing DA receptor-independent effects of DA neurons on SPNs showed a small, but rapid excitatory influence mediated by co-release of glutamate[9–12]. However, these experiments were performed in the presence of GABA receptor antagonists, precluding the detection of inhibitory influences. To observe conductances recruited after DA neuron stimulation, we performed whole-cell voltage-clamp recordings from dSPNs and iSPNs in dorsal striatum without pharmacological blockers. When SPNs were held at −70 mV (E_{Cl}, the reversal potential for chloride), nigrostriatal fibre stimulation evoked fast inward currents in approximately 75% of dSPNs ($n = 6/8$) and iSPNs ($n = 16/21$) (Fig. 2a–e). These currents showed similar properties in both cell types (Fig. 2a–c) and were consequently pooled for analysis: they exhibited peak amplitudes of 40 ± 5 pA ($n = 22$), rise times of 2.6 ± 0.4 ms and decay time constants of 6.8 ± 1.5 ms. Moreover, they reversed at approximately 0 mV and were sensitive to the α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) and NMDA (*N*-methyl-D-aspartate) receptor blockers NBQX and CPP (Fig. 2a, b, f), indicating that they are glutamatergic excitatory postsynaptic currents (EPSCs). Although glutamate release by midbrain DA neurons was proposed to be limited to mesolimbic fibres innervating ventral striatum[12], the currents reported here in dorsal striatum show similar properties[11,12], and are sensitive to the same pharmacological agents as DA release (Figs 1e and 2f), suggesting a dopaminergic origin. Moreover, EPSCs were unaffected by DA receptor antagonists (Fig. 2f), indicating that glutamate release is not secondary to DA receptor activation.

To isolate conductances that mediate SPN inhibition, the membrane potential was held at the reversal potential of ionotropic glutamate receptors (~0 mV). Under these conditions, DA neuron stimulation evoked large outward currents in all recorded dSPNs ($n = 8$) and iSPNs ($n = 21$) (Fig. 2a–c and Supplementary Fig. 3). Collectively, these currents had peak amplitudes of 617 ± 78 pA ($n = 29$; range = 0.11–1.93 nA), rise times of 2.0 ± 0.2 ms and decay time constants of 56 ± 4 ms. They reversed at E_{Cl} and were blocked by SR95531 and bicuculine, but not by the $GABA_C$ receptor antagonist TPMPA (Fig. 2a, b, g), indicating that they represent $GABA_A$ receptor-mediated inhibitory postsynaptic currents (IPSCs). Similar observations were made in mice expressing Cre under control of a tyrosine hydroxylase promoter (Supplementary Fig. 4). Consistent with previous reports[11,12], no current remained with both glutamate and $GABA_A$ receptors blocked. Thus, these data show that (1) activation of dopaminergic terminals rapidly activates ionotropic glutamate and GABA receptors in SPNs, (2) the resulting currents do not differentially affect dSPNs and iSPNs, and (3) GABAergic conductances mediate the net inhibitory effect of DA neurons on striatal output.

DA neurons may activate $GABA_A$ receptors on SPNs either by recruiting a population of GABAergic interneurons—a mechanism akin to feed-forward inhibition—or by directly releasing GABA. Several lines of evidence suggest the latter. First, the latency between light and IPSC onset averaged 2.2 ± 0.1 ms ($n = 29$; Fig. 2d, e), which may be too short to accommodate two synaptic transmission steps at 32–34 °C (ref. 16). Second, GABAergic conductances preceded or occurred synchronously with EPSCs (Fig. 2d, e), which arise after glutamate co-release from DA terminals[10–12] (Fig. 2h, i). Third, a population of dopaminergic SNc neurons expresses messenger RNA for glutamic acid decarboxylase (GAD-65)[17], indicating that they may synthesize GABA. We reasoned that if GABA originates from dopaminergic terminals, optically evoked IPSCs, EPSCs and extracellular DA should be similarly affected by pharmacological agents and persist under conditions that prevent disynaptic transmission. Indeed, IPSCs in SPNs were reduced by quinpirole, eliminated by cadmium, but were insensitive to glutamate and DA receptor inhibitors (Fig. 2g), indicating that they require Ca^{2+}-dependent release of a transmitter other than glutamate or DA. Moreover, light-evoked IPSCs, EPSCs and DA release were abolished in the presence of the voltage-gated Na^{+} channel blocker tetrodotoxin (TTX), showing that ChR2-mediated depolarization is not sufficient to trigger transmitter release from nigrostriatal axons (Fig. 2h, i). Neurotransmission can be rescued from directly illuminated ChR2-expressing terminals in the presence of TTX by providing extra depolarization with the voltage-gated K^{+} channel blocker 4-aminopyridine[18]. Accordingly, IPSCs, EPSCs and DA release recovered upon co-application of TTX and 4-aminopyridine (Fig. 2h, i), indicating that GABA, glutamate and DA are directly released from dopaminergic terminals.

The vesicular GABA transporter VGAT (encoded by *Slc32a1*) is the only transporter known to package GABA into synaptic vesicles and is considered indispensable for inhibitory synaptic transmission[19]. We generated conditional knockout (cKO) mice in which VGAT is specifically deleted from DA neurons (*Slc6a3*$^{\mathrm{IRES-Cre/wt}}$;*Slc32a1*$^{\mathrm{lox/lox}}$ mice), predicting that ChR2-evoked IPSCs would be abolished. However, light-evoked IPSCs and EPSCs were unaffected in these mice (Fig. 3a, b, g, j), indicating that VGAT is not responsible for vesicular loading of GABA in DA neurons. We instead hypothesized that GAD-65 may synthesize GABA from its metabolic precursor glutamate once inside synaptic vesicles. If correct, preventing glutamate loading into synaptic vesicles by genetically ablating the vesicular glutamate transporter 2 (VGLUT2; encoded by *Slc17a6*) from DA neurons (*Slc6a3*$^{\mathrm{IRES-Cre/wt}}$;*Slc17a6*$^{\mathrm{lox/lox}}$ mice) should eliminate IPSCs and EPSCs[10]. However, conditional deletion of VGLUT2 abolished light-evoked EPSCs in SPNs without affecting IPSC amplitude or latency (Fig. 3c, g, j), excluding this possibility.

These results indicate that GABA release originates in nigrostriatal terminals but is independent of VGAT and VGLUT2, suggesting the existence of an alternative vesicular transporter with previously unidentified activity for GABA. Consistent with this, IPSCs were eliminated in slices obtained from mice treated with the VMAT2 antagonists reserpine, Ro4-1284 or TBZ and largely recovered upon Ro4-1284 and TBZ washout (Fig. 3d–g). The same manipulation did not affect EPSCs or VGAT-dependent GABA release from SPNs (Supplementary Fig. 5), and we confirmed that DA itself does not function as a GABA$_{\mathrm{A}}$ receptor agonist in SPNs (Supplementary Fig. 6). Moreover, DA depletion with α-methyl-tyrosine (Fig. 1e) had no effect on IPSCs or EPSCs (Fig. 3g). We therefore conclude that VMAT2, but not DA, is required for the release of GABA by DA neurons.

If GABAergic IPSCs depend on VMAT2 solely for GABA transport into synaptic vesicles, IPSCs should be restored in reserpine-treated mice by expressing VGAT in DA neurons (Fig. 3h). Accordingly, in *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$ mice injected with AAV encoding Cre-dependent VGAT (AAV-DIO-VGAT) and treated with reserpine, optogenetic stimulation of nigrostriatal axons elicited large SR95531-sensitive IPSCs exhibiting synaptic latencies (Fig. 3g–j) and rise times (1.8 ± 0.3, $n = 10$; $P > 0.05$ versus control, Mann–Whitney rank sum test) indistinguishable from IPSCs observed in untreated mice. Together, these data indicate that presynaptic DA terminals contain GABA, the synaptic packaging of which requires VMAT2 but not VGAT.

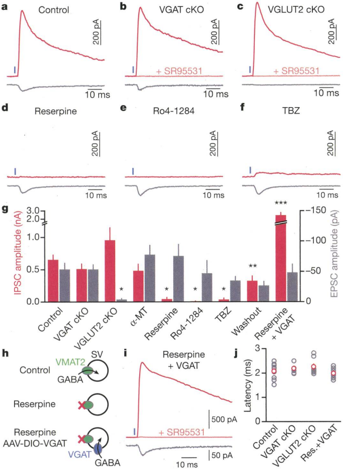


Figure 3 | GABA release from DA neurons requires VMAT2 but not VGAT. **a–c**, Representative ChR2-evoked (1 ms, blue) IPSCs (red, $V_{\mathrm{h}} = 0$ mV) and EPSCs (grey, $V_{\mathrm{h}} = -70$ mV) from SPNs in *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$ (control, **a**), *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$;*Slc32a1*$^{\mathrm{lox/lox}}$ (VGAT cKO, **b**) and *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$;*Slc17a6*$^{\mathrm{lox/lox}}$ (VGLUT2 cKO, **c**) mice. **d–e**, As in **a** for control *Slc6a3*$^{\mathrm{IRES-Cre/wt}}$ mice treated *in vivo* and *in vitro* with the VMAT2 antagonists reserpine (**d**), Ro4-1284 (**e**) or TBZ (**f**). **g**, Mean IPSC (red) and EPSC (grey) amplitudes across conditions ($n = 4$–33). Washout, slices obtained from mice treated *in vivo* with TBZ or Ro4-1284 and subsequently allowed to recover *in vitro* for more than 1 h in ACSF. $^{*}P < 0.05$ versus ACSF; $^{**}P < 0.05$ versus TBZ/Ro4-1284; $^{***}P < 0.05$ versus reserpine (Kruskal–Wallis analysis of variance). **h**, Schematic of working hypothesis: Provided DA neurons contain cytosolic GABA, viral expression of reserpine-resistant VGAT in DA neurons should rescue GABA transport into synaptic vesicles (SVs) and IPSCs in reserpine-treated mice. Note that VGAT might also incorporate into VMAT2^{-} vesicles. **i**, Voltage-clamp recording ($V_{\mathrm{h}} = 0$ mV, red; $V_{\mathrm{h}} = -70$ mV, grey) from a reserpine-treated iSPN upon optogenetic stimulation (1 ms, blue) of VGAT-expressing DA axons. **j**, IPSC onset latencies (grey, individual cells; red, average) did not differ across conditions. res., reserpine. GABAergic nature of outward currents ($V_{\mathrm{h}} = 0$ mV) in **b**, **c** and **i** was confirmed with SR95531 (pink). Error bars, s.e.m.

VMAT2 transports a variety of substrates[20], including catecholamines, serotonin and histamine. Although GABA does not bear structural resemblance to known VMAT2 substrates, our findings suggest that VMAT2 may function as a vesicular GABA transporter. To test this possibility, we asked if VMAT2 can substitute for VGAT to sustain GABA release in a population of non-monoaminergic GABAergic neurons. Specifically, we attempted to restore GABA release in iSPNs devoid of VGAT by exogenously expressing VMAT2. We conditionally excised the gene encoding VGAT in iSPNs and virally expressed ChR2 in these cells to allow monitoring of GABA release from iSPN axon collaterals onto neighbouring dSPNs as light-evoked IPSCs[21] (Fig. 4a). Whereas optogenetic stimulation of iSPNs in control mice (*Adora2a*-Cre;*Slc32a1*$^{\mathrm{lox/wt}}$;*Drd2*-EGFP) reliably evoked large IPSCs in dSPNs, IPSCs were almost entirely abolished in dSPNs from VGAT cKO mice (*Adora2a*-Cre;*Slc32a1*$^{\mathrm{lox/lox}}$;*Drd2*-EGFP) (Fig. 4b, c, e), confirming the dependence of vesicular GABA transport in SPNs

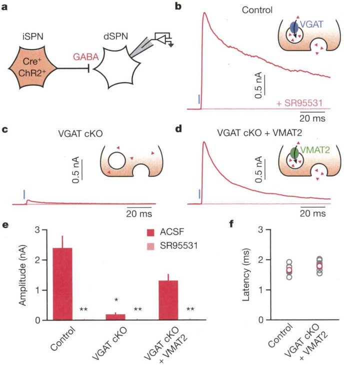


Figure 4 | VMAT2 functions as a vesicular GABA transporter. **a**, Experimental setup: ChR2 was selectively expressed in Cre-containing iSPNs of mice with one (control; *Adora2a*-Cre;*Slc32a1*$^{\text{lox/wt}}$;*Drd2*-EGFP mice) or both alleles of the gene encoding VGAT flanked by lox sites (VGAT cKO; *Adora2a*-Cre;*Slc32a1*$^{\text{lox/lox}}$;*Drd2*-EGFP mice). VGAT cKO + VMAT2, an AAV encoding Cre-dependent VMAT2 (AAV-DIO-VMAT2), was co-injected with AAV-DIO-ChR2 in the striatum of VGAT cKO mice to rescue GABA release from iSPNs. **b**–**d**, Voltage-clamp recordings ($V_h = 0$ mV) of axon-collateral IPSCs in dSPNs evoked by optogenetic stimulation (1 ms, blue) of iSPNs in the absence (red) or presence (pink) of SR95531 in control (**b**), VGAT cKO (**c**) and VGAT cKO + VMAT2 (**d**) mice. Insets: iSPN presynaptic terminal schematic illustrating experimental conditions. Red triangles, GABA. **e**, Summary histogram (mean ± s.e.m.) of experiments in **b**–**d** ($n = 10$–15 dSPNs). $*P < 0.05$ versus control and VGAT cKO+VMAT2 (Kruskal–Wallis analysis of variance); $**P < 0.05$ versus IPSC without SR95531 (Mann–Whitney rank sum test). **f**, IPSC onset latencies from light presentation onset (grey, individual cells; red, mean ± s.e.m.) in control (*Adora2a*-Cre;*Slc32a1*$^{\text{lox/wt}}$;*Drd2*-EGFP mice) and VGAT cKO+VMAT2 mice (*Adora2a*-Cre;*Slc32a1*$^{\text{lox/lox}}$;*Drd2*-EGFP mice transduced with AAV-DIO-VMAT2 in striatum).

on VGAT[19,21]. By contrast, in VGAT cKO mice transduced with Cre-dependent VMAT2 (AAV-DIO-VMAT2), light reliably evoked large, short-latency IPSCs (Fig. 4d–f). Thus, VMAT2 expression in GABAergic neurons can functionally replace VGAT and sustain inhibitory synaptic transmission.

Midbrain DA neurons are critical for the initiation, selection and reinforcement of motor actions and have been implicated in motor, cognitive and addictive disorders[1–7]. Their effects are largely attributed to the slow neuromodulatory actions of DA receptors[1–3]. Our studies show that DA neurons also exert a rapid and potent inhibitory influence on the activity of SPNs by releasing another transmitter that activates GABA$_A$ receptors. Release of this transmitter depends on VMAT2 activity, indicating that VMAT2 or a molecular complex requiring VMAT2 activity packages it in synaptic vesicles together with DA. The simplest model accounting for our data is that GABA is the transmitter packaged by VMAT2 and released onto SPNs. However, we cannot rule out that another molecule present in iSPNs and DA neurons, acting as a GABA$_A$ receptor agonist, and serving as a substrate for both VMAT2 and VGAT, fulfils that role.

It is estimated that 5–10% of SNc DA neurons express GAD-65 and fewer than 1% express VGLUT2 in rodents[10,17,22]. Therefore, distinct subpopulations of DA neurons may release GABA and glutamate, and reliable detection of IPSCs and EPSCs may result from innervation of SPNs by several DA neurons[23]. Alternatively, GABAergic transmission may be common to midbrain DA neurons as any means of acquiring GABA, such as plasma membrane uptake, would result in vesicular transport. GAD expression varies in response to neuronal activity[24] and damage[25], suggesting that GABAergic signalling from DA neurons may be altered by drug abuse. Our findings also suggest that loss of GABA release in the striatum may contribute to the symptoms of Parkinson's disease as well as to the alterations of striatal circuitry that accompany loss of midbrain DA neurons.

Because all monoaminergic neurons express VMAT2, GABA co-release may extend to adrenergic, noradrenergic, serotonergic, histaminergic as well as other dopaminergic neurons. Indeed, GABA release was reported from dopaminergic amacrine cells in the retina[26] and periglomerular cells in the olfactory bulb[27], although dependence on VMAT2 remains to be determined. Moreover, a considerable fraction (up to 80%) of monoaminergic neurons in locus coeruleus[28], hypothalamus[29] and raphe nuclei[30] contain GABA or express GAD. Together, these findings expand the repertoire of synaptic mechanisms available to monoaminergic cells and suggest that perturbations of GABA co-transmission might contribute to the aetiology of monoaminergic pathologies or to the therapeutic efficacy of VMAT2 antagonists.

METHODS SUMMARY

Stereotaxic injections of Cre-dependent AAVs were targeted to SNc of postnatal day (P)18–25 *Slc6a3*$^{\text{IRES}-\text{Cre/wt}}$;*Drd2*-EGFP mice, in which Cre and EGFP are respectively restricted to cells expressing the plasma membrane DA transporter DAT and D$_2$ receptors. Conditional deletion of VGAT or VGLUT2 in these mice was respectively achieved by breeding in *Slc32a1*$^{\text{lox}}$ or *Slc17a6*$^{\text{lox}}$ alleles to homozygosity. AAVs were allowed to express for at least 21 days before experimentation.

Whole-cell current- and voltage-clamp recordings were obtained from pathway-identified dorsal striatum SPNs in acute parasagittal slices of mature (P40–218) mice at 32–34 °C using standard techniques. Constant-potential amperometry (+400 to 600 mV versus Ag/AgCl) was performed using carbon-fibre microelectrodes. ChR2-expressing fibres were stimulated using brief full-field flashes of blue laser light (1 ms; 473 nm; 6.5–10.0 mW mm^{-2}) at intervals of 30 s or more. Traces are the average of three to five consecutive acquisitions. For pharmacological analyses, the peak amplitude of three consecutive light-evoked responses 3–4 min after drug perfusion onset were averaged, normalized to baseline and compared with values obtained at corresponding times in control preparations bathed in artificial cerebrospinal fluid (ACSF). Data in text and figures are reported as mean ± standard error of the mean (s.e.m.). Statistical tests are noted in the text (significance: $P < 0.05$).

Full Methods and any associated references are available in the online version of the paper.

Received 15 May; accepted 27 July 2012.

Published online 3 October 2012.

1. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80,** 1–27 (1998).
2. Wickens, J. R., Reynolds, J. N. & Hyland, B. I. Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* **13,** 685–690 (2003).
3. Gerfen, C. R. & Surmeier, D. J. Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34,** 441–466 (2011).
4. Palmiter, R. D. Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. *Ann. NY Acad. Sci.* **1129,** 35–46 (2008).
5. Albin, R. L., Young, A. B. & Penney, J. B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12,** 366–375 (1989).
6. Dagher, A. & Robbins, T. W. Personality, addiction, dopamine: insights from Parkinson's disease. *Neuron* **61,** 502–510 (2009).
7. Sulzer, D. How addictive drugs disrupt presynaptic dopamine neurotransmission. *Neuron* **69,** 628–649 (2011).
8. Bentivoglio, M. & Morei, M. in *Handbook of Chemical Neuroanatomy – Dopamine* Vol. 21, 1–107 (Elsevier, 2005).
9. Chuhma, N. *et al.* Dopamine neurons mediate a fast excitatory signal via their glutamatergic synapses. *J. Neurosci.* **24,** 972–981 (2004).
10. Hnasko, T. S. *et al.* Vesicular glutamate transport promotes dopamine storage and glutamate corelease *in vivo*. *Neuron* **65,** 643–656 (2010).
11. Tecuapetla, F. *et al.* Glutamatergic signaling by mesolimbic dopamine neurons in the nucleus accumbens. *J. Neurosci.* **30,** 7105–7110 (2010).

12. Stuber, G. D., Hnasko, T. S., Britt, J. P., Edwards, R. H. & Bonci, A. Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. *J. Neurosci.* **30,** 8229–8233 (2010).
13. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8,** 1263–1268 (2005).
14. Backman, C. M. *et al.* Characterization of a mouse strain expressing Cre recombinase from the 3′ untranslated region of the dopamine transporter locus. *Genesis* **44,** 383–390 (2006).
15. Schmitz, Y., Benoit-Marand, M., Gonon, F. & Sulzer, D. Presynaptic regulation of dopaminergic neurotransmission. *J. Neurochem.* **87,** 273–289 (2003).
16. Sabatini, B. L. & Regehr, W. G. Timing of synaptic transmission. *Annu. Rev. Physiol.* **61,** 521–542 (1999).
17. Gonzalez-Hernandez, T., Barroso-Chinea, P., Acevedo, A., Salido, E. & Rodriguez, M. Colocalization of tyrosine hydroxylase and GAD65 mRNA in mesostriatal neurons. *Eur. J. Neurosci.* **13,** 57–67 (2001).
18. Cruikshank, S. J., Urabe, H., Nurmikko, A. V. & Connors, B. W. Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. *Neuron* **65,** 230–245 (2010).
19. Wojcik, S. M. *et al.* A shared vesicular carrier allows synaptic corelease of GABA and glycine. *Neuron* **50,** 575–587 (2006).
20. Yelin, R. & Schuldiner, S. The pharmacological profile of the vesicular monoamine transporter resembles that of multidrug transporters. *FEBS Lett.* **377,** 201–207 (1995).
21. Kozorovitskiy, Y., Saunders, A., Johnson, C. A., Lowell, B. B. & Sabatini, B. L. Recurrent network activity drives striatal synaptogenesis. *Nature* **485,** 646–650 (2012).
22. Berube-Carriere, N. *et al.* The dual dopamine-glutamate phenotype of growing mesencephalic neurons regresses in mature rat brain. *J. Comp. Neurol.* **517,** 873–891 (2009).
23. Matsuda, W. *et al.* Single nigrostriatal dopaminergic neurons form widely spread and highly dense axonal arborizations in the neostriatum. *J. Neurosci.* **29,** 444–453 (2009).
24. Ramirez, M. & Gutierrez, R. Activity-dependent expression of GAD67 in the granule cells of the rat hippocampus. *Brain Res.* **917,** 139–146 (2001).
25. Gonzalez-Hernandez, T., Barroso-Chinea, P. & Rodriguez, M. Response of the GABAergic and dopaminergic mesostriatal projections to the lesion of the contralateral dopaminergic mesostriatal pathway in the rat. *Mov. Disord.* **19,** 1029–1042 (2004).
26. Hirasawa, H., Puopolo, M. & Raviola, E. Extrasynaptic release of GABA by retinal dopaminergic neurons. *J. Neurophysiol.* **102,** 146–158 (2009).
27. Maher, B. J. & Westbrook, G. L. Co-transmission of dopamine and GABA in periglomerular cells. *J. Neurophysiol.* **99,** 1559–1564 (2008).
28. Iijima, K. Chemocytoarchitecture of the rat locus ceruleus. *Histol. Histopathol.* **8,** 581–591 (1993).
29. Trottier, S. *et al.* Co-localization of histamine with GABA but not with galanin in the human tuberomamillary nucleus. *Brain Res.* **939,** 52–64 (2002).
30. Broadbelt, K. G., Paterson, D. S., Rivera, K. D., Trachtenberg, F. L. & Kinney, H. C. Neuroanatomic relationships between the GABAergic and serotonergic systems in the developing human medulla. *Auton. Neurosci.* **154,** 30–41 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors thank A. Saunders and Y. Kozorovitskiy for generating and characterizing the AAV-DIO-EGFP and AAV-DIO-VGAT constructs, D. Sulzer and H. Zhang for assistance with amperometry, R. Shah and C. Johnson for technical support, and members of the laboratory for discussions. This work was supported by a Nancy Lurie Marks Family Foundation postdoctoral fellowship (N.X.T.) and by grants from the National Institutes of Health (NS046579 to B.L.S. and 4R00NS075136 to J.B.D.).

Author Contributions N.X.T., J.B.D. and B.L.S. designed the experiments. N.X.T. performed the experiments described in the figures and text and analysed the data. J.B.D. performed experiments that initiated this study, devised the injection coordinates, established amperometric recordings and participated in their acquisition. N.X.T. and B.L.S. wrote the manuscript with contributions from J.B.D.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.L.S. (bsabatini@hms.harvard.edu).

doi:10.1038/nature11401

FTO genotype is associated with phenotypic variability of body mass index

A list of authors and their affiliations appears at the end of the paper.

There is evidence across several species for genetic control of phenotypic variation of complex traits[1–4], such that the variance among phenotypes is genotype dependent. Understanding genetic control of variability is important in evolutionary biology, agricultural selection programmes and human medicine, yet for complex traits, no individual genetic variants associated with variance, as opposed to the mean, have been identified. Here we perform a meta-analysis of genome-wide association studies of phenotypic variation using ~170,000 samples on height and body mass index (BMI) in human populations. We report evidence that the single nucleotide polymorphism (SNP) rs7202116 at the *FTO* gene locus, which is known to be associated with obesity (as measured by mean BMI for each rs7202116 genotype)[5–7], is also associated with phenotypic variability. We show that the results are not due to scale effects or other artefacts, and find no other experiment-wise significant evidence for effects on variability, either at loci other than *FTO* for BMI or at any locus for height. The difference in variance for BMI among individuals with opposite homozygous genotypes at the *FTO* locus is approximately 7%, corresponding to a difference of ~0.5 kilograms in the standard deviation of weight. Our results indicate that genetic variants can be discovered that are associated with variability, and that between-person variability in obesity can partly be explained by the genotype at the *FTO* locus. The results are consistent with reported *FTO* by environment interactions for BMI[8], possibly mediated by DNA methylation[9,10]. Our BMI results for other SNPs and our height results for all SNPs suggest that most genetic variants, including those that influence mean height or mean BMI, are not associated with phenotypic variance, or that their effects on variability are too small to detect even with samples sizes greater than 100,000.

Genetic studies of complex traits usually focus on quantifying and dissecting phenotypic variation within populations, by contrasting mean differences in phenotypes between genotypes. For example, in association studies the difference between the average phenotype (P) of each genotype is tested. In addition, the phenotypic variance among individuals of the same genotype (G) can vary across genotypes, so that phenotypic variance conditional on genotype, $\mathrm{var}(P|G)$, is not constant. Phenotypic variance given a particular genotype does not need to be due to sensitivity to external environmental factors but can, for example, be caused by developmental fluctuation of the internal micro-environment in a genotype-dependent manner[1]. For example, genetic control of stochastic variation in development or in homeostatic control[1,4]. The difference between genotypes can also depend on external factors, for example, on the environment in which they are reared, in which case there is a genotype by environment ($G \times E$) interaction. In species in which the same genotype can be measured across defined environments, such as in plant or animal populations, the difference in mean phenotype for each genotype can be quantified experimentally, and is known as the reaction norm of the genotype[11,12]. However, any environment is likely to be heterogeneous, so that the environment experienced by each individual differs, although these differences are not formally recognized by the experimenter. In this situation, if a $G \times E$ interaction exists it may manifest as differences in environmental sensitivity so that genotypes differ in phenotypic variance. Therefore, even if the environments, internal or external, are not directly measured, evidence for genetic control of variation can be quantified through an analysis of variability.

There is empirical evidence for genetic control of phenotypic variation in several species[1], including *Drosophila*[13], snails[14], maize[15] and chickens[3], and specific quantitative trait loci with an effect on variance have been reported for yeast[2] and *Arabidopsis*[4]. Many theories and methods to identify genetic loci responsible for phenotypic variability have been proposed[1,16–18]. In humans, there have been reports that variability of serum cholesterol and triglyceride levels within monozygotic twin pairs depends on their genotype at the MN blood group system[19]. In clinical practice, knowledge of phenotypic variability as a function of genotype may be important when the phenotypes are risk factors for disease or treatment response, in particular when there are no mean differences between genotypes in the population[19].

Detection of genetic variation in environmental or phenotypic variance requires large sample sizes because relative to their expected values, the variance has a larger sampling error than the mean[16,20]. We performed a meta-analysis of genome-wide association studies (GWAS) of phenotypic variation for height and BMI in human populations on approximately 170,000 samples comprising 133,154 in a discovery set and 36,727 for *in silico* replication, and report a single locus with a genome-wide significant effect on variability in BMI. Height and BMI were chosen because genetic effects on variability in height and size traits have been reported in other species, and because very large samples of genotyped and phenotyped individuals are available through existing research consortia.

We performed a discovery meta-analysis of 38 studies consisting of 133,154 individuals (60% females) of recent European decent to identify SNPs that are associated with the variability of height or BMI. In each study, ~2.44 million genotyped and imputed autosomal SNPs were included in the analysis after applying quality-control filters. We adjusted height and BMI phenotypes for possible covariates such as age, sex and case-control status, and standardized them to z scores by an inverse-normal transformation. We then regressed the squared z scores (z^2), which are a measure of variance[20], on the genotype indicator variable of each SNP to test for association of the SNP with trait variability. The association statistics were corrected by the genomic control method[21] in individual studies and then combined by an inverse-variance meta-analysis across all of the studies (see Methods). We selected 42 SNPs at 6 loci for height and 51 SNPs at 7 loci for BMI with $P < 5 \times 10^{-6}$ for *in silico* replication (Supplementary Fig. 1). We examined the top two SNPs at each of the 6 loci for height and 7 loci for BMI in a further sample of 36,727 individuals (54% females) of European ancestry from 13 studies (Methods). For BMI, only rs7202116 at the *FTO* locus (Fig. 1) and rs7151545 at the *RCOR1* locus (Supplementary Fig. 2) were replicated at genome-wide significance level, with $P = 2.9 \times 10^{-4}$ and $P = 3.6 \times 10^{-3}$ in the validation set and $P = 2.4 \times 10^{-10}$ and $P = 4.1 \times 10^{-8}$ in the combined set, respectively (Table 1). None of the height SNPs was replicated (Table 1). We show by an approximate conditional analysis using summary statistics from the discovery meta-analysis and estimated

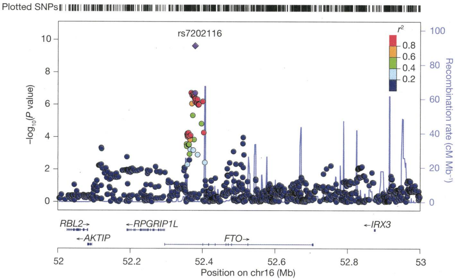


Figure 1 | Test statistics ($-\log_{10}$(*P* values)) for association with BMI variability in the discovery meta-analysis of SNPs at the *FTO* locus against their physical location. The SNPs surrounding rs7202116 are colour-coded to reflect their linkage disequilibrium with rs7202116. The recombination rates are plotted in cyan to reflect local linkage disequilibrium structure. Genes, the position of exons and the direction of transcription from the University of California, Santa Cruz (UCSC) genome browser are noted. The *P* value for rs7202116 in the combined set is represented by a purple diamond, and that from the discovery set by a purple circle.

linkage disequilibrium structure from the Atherosclerosis Risk In Communities (ARIC) cohort that there is no secondary associated SNP in the *FTO* region when conditioning on rs7202116 (Supplementary Fig. 3). The estimate of the effect associated with rs7202116 on BMI z^2 was slightly larger in men (0.041, standard error (SE) = 0.009) than in women (0.033, SE = 0.007) in the combined set but the difference was not significant ($P = 0.670$). The *RCOR1* SNP only just passed the genome-wide significance level (5×10^{-8}), however, it did not reach the experiment-wise significance level (2.5×10^{-8}) considering that two independent traits were tested. There were several case-control studies included in the meta-analysis that were ascertained for diseases that may be correlated with BMI. We performed a further meta-analysis in the combined set excluding these case-control studies, and the *FTO* SNP rs7202116 remained genome-wide significant with $P = 2.8 \times 10^{-11}$ but the *RCOR1* SNP did not with $P = 3.6 \times 10^{-5}$ (Supplementary Table 1). We therefore focus on the *FTO* locus in the main text and provide the results for the *RCOR1* locus in the Supplementary Information.

On the scale on which BMI is measured, the predicted per-allele effect of the G allele (the other allele is A) of rs7202116 on the mean difference is 0.37 kg m^{-2} in men and 0.43 kg m^{-2} in women[22], and the effect on the variance difference is 0.79 kg^2 m^{-4} in men and 1.09 kg^2 m^{-4} in women, reflecting the larger standard deviation of BMI in women compared with men (Supplementary Table 2). Assuming an additive model, the mean difference between the GG and AA genotypes is 0.74 kg m^{-2} in men and 0.86 kg m^{-2} in women, with a variance difference between the two genotypes of 1.58 kg^2 m^{-4} in men and 2.18 kg^2 m^{-4} in women, which is 7.2% of the phenotypic variance of BMI in both men and women. To provide an illustration of the effect of rs7202116 on BMI variance, we did an approximate calculation of its effect on the variance of weight. If we take the mean height of 1.78 m for men and 1.65 m for women, the difference in the variance of weight between the two genotype groups is roughly 16 kg^2 in both men and women (Supplementary Table 2). For example, if the standard deviation (SD) of weight is 15 kg for men, the predicted SD of weight in the two homozygous genotype classes is 14.73 and 15.27 kg, respectively.

The effect of a SNP on variance could be owing to our use of the z^2 value as a measure of variance or to a general relationship between mean and variance of BMI[1,23]. Below we present evidence that excludes these two explanations.

Table 1 | Associations of the top 6 and 7 loci with variance of height and BMI, respectively

Chr.	SNP	bp	Nearest gene	CA	Discovery				
					Freq.	β	SE	*P*	*n*
Height									
1	rs6429820	14,210,915	*PRDM2*	G	0.196	−0.035	0.0071	1.0×10^{-6}	129,200
2	rs6429975	143,002,110	*KYNU*	T	0.180	−0.036	0.0074	1.0×10^{-6}	129,196
2	rs6748377	45,002,877	*SIX3*	T	0.175	−0.038	0.0075	4.0×10^{-7}	129,183
7	rs10486722	41,778,433	*INHBA*	C	0.339	0.029	0.0060	1.0×10^{-6}	128,834
8	rs1026852	3,577,500	*CSMD1*	G	0.444	−0.029	0.0059	1.0×10^{-6}	126,363
14	rs12891343	34,453,301	*BAZ1A*	T	0.227	0.031	0.0068	5.0×10^{-6}	128,725
BMI									
2	rs12328474	140,638,570	*LRP1B*	G	0.263	−0.038	0.0078	1.2×10^{-6}	104,640
2	rs10932241	208,685,200	*CRYGD*	C	0.407	0.028	0.0059	2.9×10^{-6}	127,597
4	rs11942401	188,052,244	*FAT*	A	0.140	−0.043	0.0085	4.3×10^{-7}	125,010
6	rs1418304	82,795,837	*IBTK*	G	0.496	−0.026	0.0057	3.3×10^{-6}	127,611
14	rs12894649	102,232,512	*RCOR1*	C	0.057	0.061	0.0126	1.3×10^{-6}	127,080
14	**rs7151545**	**102,247,397**	***RCOR1***	**G**	**0.057**	**0.059**	**0.0126**	$\mathbf{2.4 \times 10^{-6}}$	**127,080**
16	rs7193144	52,368,187	*FTO*	C	0.403	0.030	0.0058	1.9×10^{-7}	127,537
16	**rs7202116**	**52,379,116**	***FTO***	**G**	**0.402**	**0.035**	**0.0067**	$\mathbf{2.0 \times 10^{-7}}$	**95,966**
18	rs620052	37,900,962	*PIK3C3*	G	0.378	0.033	0.0069	1.6×10^{-6}	95,971

The squared *z* scores (z^2) were used to test for association of the top 6 and 7 SNPs with trait variability (height and BMI, respectively). The discovery set consists of 133,154 individuals, and data for *in silico* replication are from another 36,727 samples. At both the *FTO* and *RCOR1* loci, the second top SNPs (highlighted in bold) in the discovery set pass the single trait genome-wide significance level (5×10^{-8}) in the combined set. β, estimate of additive effect on z^2; bp, physical position; CA, coded allele; chr., chromosome; freq., frequency of the coded allele.

If an SNP has an effect on the mean, the test statistic for association of the SNP with z^2 will be inflated, and the non-centrality parameter (NCP_{v0}) of the χ^2 test under the null hypothesis of no effect on variance is: $np(1-p)(1-2p)2(a+(1-2p)d)^4$, in which n is the sample size, p is the frequency of the coded allele, and a and d are the additive and dominance effects, respectively, on the mean difference (Supplementary Note). We show by analysis and simulation results based on an additive and dominance genetic model that such inflation is inversely proportional to the minor allele frequency (MAF) of the SNP; that is, SNPs with a lower MAF will tend to have higher test statistics under the null hypothesis (Supplementary Fig. 4). However, when we plotted the observed test statistics of the confirmed 180 height loci[24] and 32 BMI loci[22] that have the largest reported effects on the mean, we did not observe such a trend (Supplementary Fig. 5). We calculated the NCP_{v0} of the known height and BMI loci given the effects on the mean from the published papers[22,24], and the NCP_{v0} values of all these known loci were smaller than 1 (results not shown). The observed genomic inflation factor in the discovery meta-analysis was 1.039 for height and 1.033 for BMI (Supplementary Fig. 6). This small inflation could be due to many SNPs affecting the mean and therefore having a tiny effect on z^2 (Supplementary Fig. 7), or many SNPs that have an effect on the variance that is too small to be significant even with our large sample size. Across common SNPs in the genome, variants at the *FTO* locus have the largest effect size on BMI[22]. The G allele of the *FTO* SNP rs7202116 has a population frequency of ~0.4 and an additive effect on the mean BMI of ~0.1 z-score units[5,22]. If our significant result at the *FTO* locus is due only to an allelic effect on mean BMI, we would expect an allelic effect on variability of ~0.002 (predicted from the equation in the Supplementary Note), which is very small compared with the observed effect of 0.036. For some traits, the variance changes in a predictable manner as the mean changes. In this case, a scale transformation, such as a logarithmic transformation, can remove effects on the variance when they are simply due to an effect on the mean[1]. We were interested in effects of SNP on variability that would remain after a scale transformation, and therefore sought to exclude scale effects that could explain our observed association. We performed further analyses in three data sets each with approximately 20,000 individuals with individual-level genotype and phenotype data available to verify the effects of rs7202116 at the *FTO* locus on BMI variance (Methods and Table 2). We used several tests, including Bartlett's test statistic, to test for the difference in variance between the three genotypes. The Bartlett's test P value was <0.05 in each of the three data sets, regardless of whether or not the BMI phenotypes were adjusted for the mean difference, logarithm transformed or inverse-normal transformed (Table 2). In the combined analysis of the three data sets totalling 60,624 individuals, the effect of rs7202116 on the BMI z^2 score after adjusting for the mean difference was 0.030 ($P = 1.2 \times 10^{-4}$) for inverse-normal transformed BMI, 0.065 (2.3×10^{-12}) for logarithm-transformed BMI, and 0.097 (8.9×10^{-16}) for BMI without scale transformation (Table 2). The decrease of the effect of rs7202116 on BMI z^2 owing to the adjustment of the mean difference was ~0.003, in line with that of ~0.002 as predicted from the theory above. Similar conclusions as above can be drawn from the further analyses for rs7151545 at the *RCOR1* locus (Supplementary Table 3). We plotted the test statistics and estimates for the effects on the variability in our discovery meta-analysis against those for the effects on the mean from the published GIANT meta-analyses for height[24] and BMI[22], and did not find any apparent correlations except for a few outlying SNPs at the *FTO* locus (Supplementary Fig. 7). These results together suggest that the observed effect of the *FTO* SNP on variability is neither a consequence of the effect on the mean nor due to the choice of scale, and that our inverse-normal transformation is likely to be overly conservative. Results from reported quantile regression of untransformed BMI on a multiple SNP predictor of BMI and on *FTO*[25] are consistent with our results but are also consistent with scale effects due to the skewed distribution of untransformed BMI. We have shown in this study that the effect of *FTO* on variability is not due to a scale effect and, concordantly, a quantile regression of both transformed and untransformed BMI z-scores on the SNPs at the *FTO* and *RCOR1* loci on BMI on 17,974 individuals shows a relationship between effect size and the quantile of the distribution (Supplementary Fig. 8). By contrast, the use of untransformed BMI induces widespread correlation between estimated SNP effects on the mean and on variance (Supplementary Fig. 9).

We have reported a meta-analysis of GWAS of squared normalized residuals for two quantitative traits in human populations, and provide empirical evidence that the *FTO* and *RCOR1* loci influence phenotypic variance of obesity. Conversely, we did not observe any significant SNPs for height or any significant SNPs other than those at the *FTO* and *RCOR1* loci for BMI to be genome-wide significantly associated with phenotypic variance (Table 1), even for those loci known to have effects on the mean (Supplementary Fig. 5), which indicates that SNP effects on variance are uncommon for height and BMI, and those previously identified SNP effects on the mean, although very small, are robust to environmental perturbation. We provide evidence that the association between the *FTO* locus and BMI variability is not due to artefacts such as scale or ascertainment. We also discuss that it is implausible that the observed effect of the *FTO* SNP on variance is due to its strong linkage disequilibrium ($D' = 1$) with a causal variant that has a large effect on the mean (Supplementary Note). The *FTO*

Table 1 | Continued

In silico replication					Combined			
Freq.	β	SE	P	n	β	SE	P	n
0.209	−0.002	0.0131	8.9×10^{-1}	32,355	−0.027	0.0062	1.0×10^{-5}	161,555
0.177	−0.002	0.0137	8.9×10^{-1}	32,472	−0.028	0.0065	1.0×10^{-5}	161,668
0.185	−0.006	0.0138	6.7×10^{-1}	31,988	−0.031	0.0066	3.0×10^{-6}	161,171
0.318	−0.005	0.0112	6.3×10^{-1}	32,416	0.021	0.0053	6.0×10^{-5}	161,250
0.435	−0.004	0.0110	7.4×10^{-1}	31,837	−0.023	0.0052	7.0×10^{-6}	158,200
0.225	0.012	0.0120	3.2×10^{-1}	36,150	0.027	0.0059	6.0×10^{-6}	164,875
0.250	0.035	0.0152	2.0×10^{-2}	32,403	−0.023	0.0069	1.1×10^{-3}	137,043
0.411	−0.006	0.0125	6.2×10^{-1}	28,641	0.022	0.0053	5.6×10^{-5}	156,238
0.128	0.003	0.0187	8.5×10^{-1}	28,016	−0.035	0.0077	6.2×10^{-6}	153,026
0.493	0.004	0.0103	6.9×10^{-1}	36,721	−0.019	0.0050	1.2×10^{-4}	164,332
0.050	0.058	0.0248	1.9×10^{-2}	32,298	0.060	0.0112	7.9×10^{-8}	159,378
0.053	**0.083**	**0.0285**	$\mathbf{3.6 \times 10^{-3}}$	**28,040**	**0.063**	**0.0115**	$\mathbf{4.1 \times 10^{-8}}$	**155,120**
0.406	0.020	0.0115	8.0×10^{-2}	32,449	0.028	0.0052	5.4×10^{-8}	159,986
0.417	**0.039**	**0.0107**	$\mathbf{2.9 \times 10^{-4}}$	**35,267**	**0.036**	**0.0057**	$\mathbf{2.4 \times 10^{-10}}$	**131,233**
0.382	−0.010	0.0111	3.7×10^{-1}	34,668	0.021	0.0059	3.5×10^{-4}	130,639

Table 2 | Effects of the *FTO* SNP rs7202116 on BMI

	BMI		log(BMI)		BMI (inv. norm.)	
	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.
WGHS (*n* = 22,888)						
β	0.148	0.142	0.100	0.093	0.046	0.040
SE	0.021	0.020	0.015	0.015	0.013	0.013
P	4.5×10^{-13}	4.0×10^{-12}	5.5×10^{-11}	8.6×10^{-10}	6.8×10^{-4}	3.3×10^{-3}
Permutation *P*	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$	9.0×10^{-4}	3.9×10^{-3}
Bartlett's *P*	1.1×10^{-24}	1.1×10^{-24}	2.0×10^{-11}	2.0×10^{-11}	6.5×10^{-3}	6.6×10^{-3}
Mean AA	−0.070	0.0	−0.069	0.0	−0.068	0.0
Mean AG	−0.001	0.0	−0.001	0.0	0.0	0.0
Mean GG	0.161	0.0	0.159	0.0	0.152	0.0
Variance AA	0.895	0.900	0.932	0.937	0.971	0.977
Variance AG	1.002	1.008	0.995	1.001	0.990	0.996
Variance GG	1.194	1.202	1.132	1.138	1.060	1.066
EPIC (*n* = 19,762)						
β	0.077	0.076	0.049	0.048	0.027	0.026
SE	0.021	0.021	0.017	0.017	0.014	0.014
P	1.7×10^{-4}	2.1×10^{-4}	3.2×10^{-3}	3.9×10^{-3}	6.1×10^{-2}	7.1×10^{-2}
Permutation *P*	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$	4.9×10^{-3}	5.1×10^{-3}	6.4×10^{-2}	7.1×10^{-2}
Bartlett's *P*	7.6×10^{-7}	7.6×10^{-7}	3.0×10^{-3}	3.0×10^{-3}	1.2×10^{-1}	1.2×10^{-1}
Mean AA	−0.077	0.000	−0.076	0.000	−0.075	0.000
Mean AG	0.012	0.000	0.012	0.000	0.012	0.000
Mean GG	0.103	0.000	0.102	0.000	0.100	0.000
Variance AA	0.932	0.936	0.951	0.955	0.967	0.970
Variance AG	1.005	1.009	1.007	1.011	1.010	1.013
Variance GG	1.085	1.089	1.045	1.049	1.013	1.017
ARIC + QIMR + NHS + HPFS (*n* = 17,974)						
β	0.070	0.067	0.049	0.046	0.026	0.024
SE	0.022	0.022	0.017	0.017	0.015	0.015
P	1.7×10^{-3}	2.8×10^{-3}	3.6×10^{-3}	6.1×10^{-3}	8.9×10^{-2}	1.2×10^{-1}
Permutation *P*	1.6×10^{-3}	2.6×10^{-3}	3.8×10^{-3}	7.1×10^{-3}	8.7×10^{-2}	1.2×10^{-1}
Bartlett's *P*	1.2×10^{-7}	1.2×10^{-7}	2.5×10^{-4}	2.5×10^{-4}	2.0×10^{-2}	2.0×10^{-2}
Mean AA	−0.067	0.0	−0.068	0.0	−0.069	0.0
Mean AG	0.006	0.0	0.008	0.0	0.010	0.0
Mean GG	0.122	0.0	0.118	0.0	0.113	0.0
Variance AA	0.968	0.973	0.978	0.983	0.994	0.998
Variance AG	0.968	0.972	0.974	0.978	0.975	0.979
Variance GG	1.131	1.136	1.093	1.097	1.059	1.064
Combined (*n* = 60,624)						
β	0.100	0.097	0.068	0.065	0.034	0.030
SE	0.012	0.012	0.009	0.009	0.008	0.008
P	8.9×10^{-17}	8.9×10^{-16}	1.4×10^{-13}	2.3×10^{-12}	2.4×10^{-5}	1.2×10^{-4}
Bartlett's *P*	1.3×10^{-32}	1.3×10^{-32}	8.5×10^{-15}	8.6×10^{-15}	4.4×10^{-4}	4.2×10^{-4}
Mean AA	−0.071	0.0	−0.071	0.0	−0.070	0.0
Mean AG	0.005	0.0	0.006	0.0	0.007	0.0
Mean GG	0.129	0.0	0.127	0.0	0.122	0.0
Variance AA	0.93	0.93	0.95	0.96	0.98	0.98
Variance AG	0.99	1.00	0.99	1.00	0.99	1.00
Variance GG	1.14	1.14	1.09	1.09	1.04	1.05

The effects of the *FTO* SNP rs7202116 on the variance for BMI and log(BMI) were tested in three subsets of data. The BMI phenotypes were corrected for age effect and standardized to *z* scores using the mean and standard deviation, or by an inverse-normal (inv. norm.) transformation in each gender group in each cohort. Phenotypes were adjusted (adj.) (or unadjusted (unadj.)) for mean difference in the three genotypes. For the EPIC cohort, 2,397 samples were in the meta-analysis, and 17,376 were not part of the meta-analysis. For the combined ARIC, QIMR, NHS and HPFS cohort, 12,741 samples were in the meta-analysis and 5,233 samples were not. β, the effect of the G allele on z^2; Bartlett's *P*, *P* value calculated from the Bartlett's test for variance difference in the three genotypes; EPIC, European Prospective Investigation into Cancer; HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; permutation *P*, empirical *P* value calculated from 10,000 permutations; QIMR, Queensland Institute of Medical Research; WGHS, Women's Genome Health Study.

SNPs that are associated with variance are also associated with mean differences in BMI. Interestingly, this phenomenon seems to be restricted to the *FTO* gene and to obesity, because we did not observe such effects for height or for BMI at loci other than *FTO*. One possible explanation of the observation is a differential response to physical activity[26], because interactions between *FTO* genotypes and physical activity have been reported for the same SNPs as we report in this study: the G allele that is associated with an increase in mean BMI has a smaller effect in the group of people with a high level of physical activity than in the absence of physical activity[8,27,28]. There may be other unknown lifestyle factors, including diet, that also interact with the *FTO* genotype and result in the observed effect on variability.

We do not provide a mechanism of how alleles at *FTO* influence variability (how *FTO* alleles affect the mean is also not known). However, the fact that the allele that increases obesity also increases variability suggests a breakdown of homeostatic control. Data on mice lacking the *Fto* gene suggest that the observed effects on mean obesity in humans may be due to upregulation or dysregulation of *FTO* expression, resulting in an increased susceptibility to obesity[29]. If both upregulation and impairment of *FTO* expression have a role then this could provide a mechanism of the observed effect on variability. The FTO protein affects demethylation of nuclear RNA *in vitro*[29], but whether the efficiency of this process depends on the *FTO* genotype or how this may be related to the observed effects on BMI is not clear. Notably, a recent study reported that rs7202116 allele G, which is present on the obesity-susceptibility haplotype at the *FTO* locus, creates a CpG site along with other variants in perfect linkage disequilibrium with it[9], and therefore risk alleles have increased DNA methylation. In addition, it was reported that a CpG site in the first intron of *FTO* showed significant hypomethylation in type 2 diabetes cases relative to controls[30], and that the risk variant seems to have an effect on methylation status at other genes[10]. DNA methylation can be affected by environmental influences, including dietary and lifestyle factors, and may affect gene expression. For example, physical exercise may increase gene expression at the *FTO* locus, but less so in GG individuals compared with AA individuals because their alleles are more methylated. This therefore suggests a possible mechanism for the observed effects on both the mean and variability. However, more research is needed to determine the molecular effect and mechanism of *FTO* on both the levels and variability of obesity.

Overall, our findings are consistent with a low heritability of phenotypic variability[1] and no common genetic variants that account for a large proportion of variation in environmental or phenotypic variability. They also indicate an absence of widespread genotype-by-environment interaction effects, at least for height and obesity in humans and with interaction effects large enough to be detected in our study in which specific environmental factors were not identified. Nevertheless, the demonstration that individual genetic loci with effects on variability can be identified with sufficiently large sample sizes facilitates further study to understand the function and evolution of the genetic control of variation.

METHODS SUMMARY

We performed a meta-analysis of 51 GWAS with 169,881 individuals of European ancestry, and $\sim$2.44 million genotyped or imputed SNPs after quality control. In each study, association analysis of each SNP with height and BMI z^2 was performed after adjustment for covariates and followed by an inverse-normal transformation. We meta-analysed the association results of each SNP from 38 studies with 133,154 individuals as a discovery set, and validated the top SNPs identified in the discovery set with association P values $< 5 \times 10^{-6}$ in a separate sample of 36,727 individuals from 13 studies. Further analyses using individual-level genotype and phenotype data to test for difference in variance of BMI between the three groups for the top SNPs at the *FTO* and *RCOR1* loci were performed on 60,624 individuals, including 22,598 individuals who were not part of the meta-analysis.

Full Methods and any associated references are available in the online version of the paper.

Received 15 January; accepted 6 July 2012.

Published online 16 September 2012.

1. Hill, W. G. & Mulder, H. A. Genetic analysis of environmental variation. *Genet Res.* **92,** 381–395 (2010).
2. Ansel, J. *et al.* Cell-to-cell stochastic variation in gene expression is a complex genetic trait. *PLoS Genet.* **4,** e1000049 (2008).
3. Wolc, A., White, I. M., Avendano, S. & Hill, W. G. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. *Poult. Sci.* **88,** 1156–1161 (2009).
4. Jimenez-Gomez, J. M., Corwin, J. A., Joseph, B., Maloof, J. N. & Kliebenstein, D. J. Genomic analysis of QTLs and genes altering natural variation in stochastic noise. *PLoS Genet.* **7,** e1002295 (2011).
5. Frayling, T. M. *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316,** 889–894 (2007).
6. Dina, C. *et al.* Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature Genet.* **39,** 724–726 (2007).
7. Scuteri, A. *et al.* Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3,** e115 (2007).
8. Kilpeläinen, T. O. *et al.* Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med.* **8,** e1001116 (2011).
9. Bell, C. G. *et al.* Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS ONE* **5,** e14040 (2010).
10. Almén, M. S. *et al.* Genome wide analysis reveals association of a *FTO* gene variant with epigenetic changes. *Genomics* **99,** 132–137 (2012).
11. Falconer, D. S. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genet. Res.* **56,** 57–70 (1990).
12. Jinks, J. L. & Connolly, V. Selection for specific and general response to environmental differences. *Heredity* **30,** 33–40 (1973).
13. Mackay, T. F. & Lyman, R. F. *Drosophila* bristles and the nature of quantitative genetic variation. *Phil. Trans. R. Soc. Lond. B* **360,** 1513–1527 (2005).
14. Ros, M. *et al.* Evidence for genetic control of adult weight plasticity in the snail Helix aspersa. *Genetics* **168,** 2089–2097 (2004).
15. Ordas, B., Malvar, R. A. & Hill, W. G. Genetic variation and quantitative trait loci associated with developmental stability and the environmental correlation between traits in maize. *Genet Res.* **90,** 385–395 (2008).
16. Yang, Y., Christensen, O. F. & Sorensen, D. Use of genomic models to study genetic control of environmental variance. *Genet. Res.* **93,** 125–138 (2011).
17. Rönnegård, L. & Valdar, W. Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics* **188,** 435–447 (2011).
18. Paré, G., Cook, N. R., Ridker, P. M. & Chasman, D. I. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS Genet.* **6,** e1000981 (2010).
19. Martin, N. G., Rowell, D. M. & Whitfield, J. B. Do the MN and Jk systems influence environmental variability in serum lipid levels? *Clin. Genet.* **24,** 1–14 (1983).
20. Visscher, P. M. & Posthuma, D. Statistical power to detect genetic loci affecting environmental sensitivity. *Behav. Genet.* **40,** 728–733 (2010).
21. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55,** 997–1004 (1999).
22. Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genet.* **42,** 937–948 (2010).
23. Struchalin, M. V., Dehghan, A., Witteman, J. C., van Duijn, C. & Aulchenko, Y. S. Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. *BMC Genet.* **11,** 92 (2010).
24. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467,** 832–838 (2010).
25. Williams, P. T. Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height. *PLoS ONE* **7,** e28764 (2012).
26. Silventoinen, K. *et al.* Modification effects of physical activity and protein intake on heritability of body size and composition. *Am. J. Clin. Nutr.* **90,** 1096–1103 (2009).
27. Andreasen, C. H. *et al.* Low physical activity accentuates the effect of the *FTO* rs9939609 polymorphism on body fat accumulation. *Diabetes* **57,** 95–101 (2008).
28. Rampersaud, E. *et al.* Physical activity and the association of common *FTO* gene variants with body mass index and obesity. *Arch. Intern. Med.* **168,** 1791–1797 (2008).
29. Jia, G. *et al. N*6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chem. Biol.* **7,** 885–887 (2011).
30. Toperoff, G. *et al.* Genome-wide survey reveals predisposing diabetes type 2-related DNA methylation variations in human peripheral blood. *Hum. Mol. Genet.* **21,** 371–383 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge funding from the Australian National Health and Medical Research Council (NHMRC grants 241944, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 496688, 552485, 613672, 613601 and 1011506), the US National Institutes of Health (grants AA07535, AA10248, AA014041, AA13320, AA13321, AA13326, DA12854 and GM057091) and the Australian Research Council (ARC grant DP1093502). A detailed list of acknowledgements by study is provided in the Supplementary Information. We apologize to authors whose work we could not cite owing to space restrictions.

Author Contributions P.M.V., M.E.G. and J.Y. conceived and designed the study. J.Y. and P.M.V. derived the analytical theory. J.Y. performed the meta-analyses and simulations. J.Y. and P.M.V. wrote the first draft of the manuscript. J.Y., D.I.C., J.H.Z. and R.J.F.L. performed further statistical verification analyses. D.P.S., W.G.H., R.J.F.L., S.I.B. and H. Snieder contributed important additional concepts and critically reviewed the manuscript before submission. S.E.M., P.A.F.M., A.C.H., N.G.M., D.R.N. and G.W.M. contributed the individual-level genotype and phenotype data of the QIMR cohort. T.M.F., J.N.H. and R.J.F.L. liaised with the GIANT consortium for this project. The cohort-specific contributions of all other authors are provided in the Supplementary Information.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.M.V. (peter.visscher@uq.edu.au).

Jian Yang[1,2], Ruth J. F. Loos[3,4], Joseph E. Powell[1,2], Sarah E. Medland[2], Elizabeth K. Speliotes[5,6], Daniel I. Chasman[7,8], Lynda M. Rose[7], Gudmar Thorleifsson[9], Valgerdur Steinthorsdottir[9], Reedik Mägi[10,11], Lindsay Waite[12], Albert Vernon Smith[13,14], Laura M. Yerges-Armstrong[15], Keri L. Monda[16], David Hadley[17], Anubha Mahajan[11], Guo Li[18], Karen Kapur[19,20], Veronique Vitart[21], Jennifer E. Huffman[21], Sophie R. Wang[22,23,24], Cameron Palmer[23,24], Tõnu Esko[10], Krista Fischer[10], Jing Hua Zhao[3], Ayşe Demirkan[25], Aaron Isaacs[25], Mary F. Feitosa[26], Jian'an Luan[3], Nancy L. Heard-Costa[27], Charles White[27], Anne U. Jackson[28], Michael Preuss[29,30], Andreas Ziegler[30], Joel Eriksson[31], Zoltán Kutalik[19,20], Francesca Frau[32], Ilja M. Nolte[33], Jana V. Van Vliet-Ostaptchouk[34,35], Jouke-Jan Hottenga[36], Kevin B. Jacobs[37], Niek Verweij[38], Anuj Goel[11,39], Carolina Medina-Gomez[40,41,42], Karol Estrada[40,41,42], Jennifer Lynn Bragg-Gresham[43], Serena Sanna[44], Carlo Sidore[43,45], Jonathan Tyrer[46], Alexander Teumer[47], Inga Prokopenko[11,48], Massimo Mangino[49], Cecilia M. Lindgren[11], Themistocles L. Assimes[50], Alan R. Shuldiner[15,51], Jennie Hui[52,53,54], John P. Beilby[52,53], Wendy L. McArdle[55], Per Hall[56], Talin Haritunians[57], Lina Zgaga[58,59], Ivana Kolcic[60], Ozren Polasek[60], Tatijana Zemunik[60], Ben A. Oostra[25], M. Juhani Junttila[61], Henrik Grönberg[56], Stefan Schreiber[62], Annette Peters[63,64], Andrew A. Hicks[65], Jonathan Stephens[66,67], Nicola S. Foad[66,67], Jaana Laitinen[68], Anneli Pouta[69,70], Marika Kaakinen[71], Gonneke Willemsen[36], Jacqueline M. Vink[36], Sarah H. Wild[58], Gerjan Navis[72], Folkert W. Asselbergs[73], Georg Homuth[47], Ulrich John[74], Carlos Iribarren[75], Tamara Harris[76], Lenore Launer[76], Vilmundur Gudnason[13,14], Jeffrey R. O'Connell[15], Eric Boerwinkle[77], Gemma Cadby[78], Lyle J. Palmer[78], Alan L. James[79,80], Arthur W. Musk[79,81], Erik Ingelsson[56], Bruce M. Psaty[82,83], Jacques S. Beckmann[19,84], Gerard Waeber[85], Peter Vollenweider[85], Caroline Hayward[21], Alan F. Wright[21], Igor Rudan[58,60], Leif C. Groop[86], Andres Metspalu[10], Kay Tee Khaw[87], Cornelia M. van Duijn[25], Ingrid B. Borecki[26,88], Michael A. Province[26,88], Nicholas J. Wareham[3], Jean-Claude Tardif[89,90], Heikki V. Huikuri[61], L. Adrienne Cupples[27,91], Larry D. Atwood[27], Caroline S. Fox[91], Michael Boehnke[28], Francis S. Collins[92], Karen L. Mohlke[93], Jeanette Erdmann[29,94], Heribert Schunkert[29,94], Christian Hengstenberg[95], Klaus Stark[95], Mattias Lorentzon[31], Claes Ohlsson[31], Daniele Cusi[32], Jan A. Staessen[96,97], Melanie M. Van der Klauw[34,35], Peter P. Pramstaller[98,99,100], Sekar Kathiresan[91,101,102,103,104], Jennifer D. Jolley[66,67], Samuli Ripatti[105,106,107], Marjo-Riitta Jarvelin[69,71,108], Eco J. C. de Geus[36], Dorret I. Boomsma[36], Brenda Penninx[109], James F. Wilson[58], Harry Campbell[58], Stephen J. Chanock[110], Pim van der Harst[38], Anders Hamsten[111,112], Hugh Watkins[11,39], Albert Hofman[41,42], Jacqueline C. Witteman[41,42], M. Carola Zillikens[40], André G. Uitterlinden[40,41,42], Fernando Rivadeneira[40,41,42], M. Carola Zillikens[40], Lambertus A. Kiemeney[113], Sita H. Vermeulen[113], Goncalo R. Abecasis[43], David Schlessinger[114], Sabine Schipf[115], Michael Stumvoll[116,117], Anke Tönjes[116,117], Tim D. Spector[49], Kari E. North[118], Guillaume Lettre[89,90], Mark I. McCarthy[11,48,119], Sonja I. Berndt[110], Andrew C. Heath[120], Pamela A. F. Madden[120], Dale R. Nyholt[2], Grant W. Montgomery[2], Nicholas G. Martin[2], Barbara McKnight[121], David P. Strachan[17], William G. Hill[122], Harold Snieder[33,35], Paul M. Ridker[7,8], Unnur Thorsteinsdottir[9,123], Kari Stefansson[9,123], Timothy M. Frayling[124], Joel N. Hirschhorn[22,23,24], Michael E. Goddard[125,126] & Peter M. Visscher[1,2,127]

[1]University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4102, Australia. [2]Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia. [3]MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge CB2 0QQ, UK. [4]Mount Sinai School of Medicine, New York, New York 10029, USA. [5]Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan 48109, USA. [6]Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan 48109, USA. [7]Division of Preventive Medicine, Brigham and Women's Hospital, 900 Commonwealth Avenue, Boston, Massachusetts 02215, USA. [8]Harvard Medical School, Boston, Massachusetts 02215, USA. [9]deCODE genetics, IS-101 Reykjavik, Iceland. [10]Estonian Genome Center, University of Tartu, Tartu 50410, Estonia. [11]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. [12]Hudson Alpha Institute for Biotechnology, Huntsville, Alabama 35806, USA. [13]Icelandic Heart Association, IS-201 Kopavogur, Iceland. [14]University of Iceland, IS-101 Reykjavik, Iceland. [15]Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA. [16]Department of Epidemiology, The University of North Carolina, Chapel Hill, North Carolina 27514, USA. [17]Division of Population Health Sciences & Education, St George's, University of London, London SW17 0RE, UK. [18]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington 98101, USA. [19]Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland. [20]Swiss Institute of Bioinformatics, 1005 Lausanne, Switzerland. [21]MRC HGU at the MRC IGMM at the University of Edinburgh, Edinburgh EH8 9AG, UK. [22]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [23]Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital, Boston, Massachusetts 02115, USA. [24]Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA. [25]Department of Epidemiology, Subdivison Genetic Epidemiology, Erasmus MC, Rotterdam, The Netherlands. [26]Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA. [27]Boston University, Boston, Massachusetts 02118, USA. [28]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA. [29]Universität zu Lübeck, Medizinische Klinik II, Ratzeburger Allee 160, 23538 Lübeck, Germany. [30]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, 23562 Lübeck, Germany. [31]Center for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 413 45 Gothenburg, Sweden. [32]Department of Health Sciences, University of Milano, 20133 Milano, Italy. [33]Unit of Genetic Epidemiology & Bioinformatics, Department of Epidemiology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [34]Department of Endocrinology, University Medical Center Groningen, University of Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands. [35]LifeLines Cohort Study, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [36]Department of Biological Psychology, VU University, 1081 BT Amsterdam, The Netherlands. [37]Core Genotyping Facility, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702, USA. [38]Department of Cardiology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [39]Cardiovascular Medicine, University of Oxford, Oxford OX3 7BN, UK. [40]Department of Internal Medicine, Erasmus MC, Rotterdam 3015GE, The Netherlands. [41]Department of Epidemiology, Erasmus MC, Rotterdam 3015GE, The Netherlands. [42]Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), 2300 RC Leiden, The Netherlands. [43]Biostatistics - Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA. [44]Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato 09042, Italy. [45]Dipartimento di Scienze Biomediche, Università di Sassari, 07100 SS, Italy. [46]Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK. [47]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17487 Greifswald, Germany. [48]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford OX3 7BN, UK. [49]Department of Twin Research and Genetic Epidemiology, King's College London, Lambeth Palace Road, London SE1 7EH, UK. [50]Department of Medicine, Stanford University School of Medicine, Stanford 94305, California, USA. [51]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland 21201, USA. [52]Pathology and Laboratory Medicine, University of Western Australia, Nedlands Western Australia 6009, Australia. [53]Molecular Genetics, PathWest Laboratory Medicine WA, University of Western Australia, Nedlands Western Australia 6009, Australia. [54]School of Population Health, University of Western Australia, Nedlands Western Australia 6009, Australia. [55]School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK. [56]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, SE-171 77 Stockholm, Sweden. [57]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. [58]Centre for Population Health Sciences, The University of Edinburgh Medical School, Edinburgh EH16 4TJ, UK. [59]Andrija Stampar School of Public Health, Medical School University of Zagreb, Zagreb, Croatia. [60]Faculty of Medicine, University of Split, Soltanska 2, 21000 Split, Croatia. [61]Institute of Clinical Medicine, Department Of Internal Medicine, University of Oulu, 90014 Oulu, Finland. [62]Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, 24098 Kiel, Germany. [63]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. [64]Munich Heart Alliance, 80802 Munich, Germany. [65]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), 39100 Bolzano, Italy. [66]Department of Haematology, University of Cambridge, Cambridge CB2 0PT, UK. [67]NHS Blood and Transplant, Cambridge, Cambridge CB2 0PT, UK. [68]Finnish Institute of Occupational Health, 90220 Oulu, Finland. [69]National Institute for Health and Welfare, 90101 Oulu, Finland. [70]Department of Clinical Sciences/Obstetrics and Gynecology, University of Oulu, 90014 Oulu, Finland. [71]Institute of Health Sciences, Biocenter Oulu, University of Oulu, 90014 Oulu, Finland. [72]Department of Internal Medicine, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [73]Department of Cardiology, Division Heart & Lungs, University medical Center Utrecht, 3508 GA Utrecht, The Netherlands. [74]Institute of Epidemiology and Social Medicine, University Medicine Greifswald, 17475 Greifswald, Germany. [75]Division of Research, Kaiser Permanente Northern California, Oakland, California 94612, USA. [76]National Institutes on Aging, National Institutes of Health, Bethesda, Maryland 20892, USA. [77]Human Genetics Center and Division of Epidemiology, The University of Texas Health Science Centers, Houston, Texas 77030, USA. [78]Genetic Epidemiology and Biostatistics Platform, Ontario Institute for Cancer Research, Toronto, Ontario M5G 1L7, Canada. [79]Medicine and Pharmacology, University of Western Australia, Nedlands Western Australia 6009, Australia. [80]Pulmonary Physiology, Sir Charles Gairdner Hospital, University of Western Australia, Nedlands Western Australia 6009, Australia. [81]Respiratory Medicine, Sir Charles Gairdner Hospital, University of Western Australia, Nedlands Western Australia 6009, Australia. [82]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, Washington 98101, USA. [83]Group Health Research Institute, Group Health Cooperative, Seattle, Washington 98101, USA. [84]Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, 1011 Lausanne, Switzerland. [85]Department of Internal Medicine, University Hospital, 1011 Lausanne, Switzerland. [86]Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden. [87]Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. [88]Division of Biostatistics, Washington University School of Medicine, St Louis, Missouri 63110, USA. [89]Département de Médecine, Université de Montréal, Montréal, Québec H4J 1C5, Canada. [90]Montreal Heart Institute, Montréal, Québec H1T 1C8, Canada. [91]Framingham Heart Study of the National Heart, Lung, and Blood Institute and Boston University, Framingham, Massachusetts 01702, USA. [92]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. [93]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599-7264, USA. [94]Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK), Universität zu Lübeck, 23562 Lübeck, Germany. [95]Klinik und Poliklinik für Innere Medizin II, 93053 Regensburg, Germany. [96]Department of Cardiovascular Diseases, University of Leuven, 3000 Leuven, Belgium. [97]Department of Epidemiology, Maastricht University, 6200 MD Maastricht, The Netherlands. [98]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), 39100 Bolzano, Italy. [99]Department of Neurology, General Central Hospital, 39100 Bolzano, Italy. [100]Department of Neurology, University of Lübeck, 23562 Lübeck, Germany. [101]Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [102]Center for Human Genetics Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [103]Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [104]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. [105]Institute for Molecular Medicine Finland, FIMM, University of Helsinki, 00014 Helsinki, Finland. [106]Public Health Genomics Unit, National Institute for Health and Welfare, 00271 Helsinki, Finland. [107]Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. [108]Department of Epidemiology and Biostatistics, MRC-HPA Center for Environment and Health, Imperial College London, London W2 1PG, UK. [109]Department of Psychiatry, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands. [110]Division of Cancer Epidemiology & Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20852, USA. [111]Karolinska Institutet, 171 77 Stockholm, Sweden. [112]Atherosclerosis Research Unit, Department of Medicine, Solna, 171 76 Stockholm, Sweden. [113]Epidemiology, Biostatistics & HTA, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands. [114]National Institute on Aging, National Institutes of Health, Bethesda, Maryland 20892, USA. [115]Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald, Germany. [116]Department of Medicine, University of Leipzig, 04103 Leipzig, Germany. [117]IFB Adiposity Diseases, University of Leipzig, 04103 Leipzig, Germany. [118]Department of Epidemiology and Carolina Center for Genome Sciences, The University of North Carolina, Chapel Hill, North Carolina 27514, USA. [119]Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Old Road Headington, Oxford OX3 7LJ, UK. [120]Department of Psychiatry, Washington University St Louis, Missouri 63110, USA. [121]Department of Biostatistics, University of Washington, Seattle, Washington 98115, USA. [122]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK. [123]Faculty of Medicine, University of Iceland, IS-101 Reykjavik, Iceland. [124]Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter EX1 2LU, UK. [125]Department of Food and Agricultural Systems, University of Melbourne, Victoria 3010, Australia. [126]Biosciences Research Division, Department of Primary Industries, Bundoora, Victoria 3083, Australia. [127]Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia.

doi:10.1038/nature11431

Controlling interneuron activity in *Caenorhabditis elegans* to evoke chemotactic behaviour

Askin Kocabas[1,2]*, Ching-Han Shen[1,2]*, Zengcai V. Guo[3] & Sharad Ramanathan[1,2,4,5,6]

Animals locate and track chemoattractive gradients in the environment to find food. With its small nervous system, *Caenorhabditis elegans* is a good model system[1,2] in which to understand how the dynamics of neural activity control this search behaviour. Extensive work on the nematode has identified the neurons that are necessary for the different locomotory behaviours underlying chemotaxis through the use of laser ablation[3–7], activity recording in immobilized animals and the study of mutants[4,5]. However, we do not know the neural activity patterns in *C. elegans* that are sufficient to control its complex chemotactic behaviour. To understand how the activity in its interneurons coordinate different motor programs to lead the animal to food, here we used optogenetics and new optical tools to manipulate neural activity directly in freely moving animals to evoke chemotactic behaviour. By deducing the classes of activity patterns triggered during chemotaxis and exciting individual neurons with these patterns, we identified interneurons that control the essential locomotory programs for this behaviour. Notably, we discovered that controlling the dynamics of activity in just one interneuron pair (AIY) was sufficient to force the animal to locate, turn towards and track virtual light gradients. Two distinct activity patterns triggered in AIY as the animal moved through the gradient controlled reversals and gradual turns to drive chemotactic behaviour. Because AIY neurons are post-synaptic to most chemosensory and thermosensory neurons[8], it is probable that these activity patterns in AIY have an important role in controlling and coordinating different taxis behaviours of the animal.

Organisms, from bacteria to multicellular eukaryotes, have to search for food to survive. Complex internal circuits process external signals to evoke and coordinate multiple motor programs, leading the animal to track attractive odours and find food. Whether there are master nodes in the circuits that control and coordinate search behaviour, and whether the neural circuits generating chemotactic behaviour in *C. elegans* can be controlled through such key nodes are important questions.

The nematode *C. elegans* uses reversals (backward movement) and sharp and gradual turns to locate and track gradients of chemoattractive

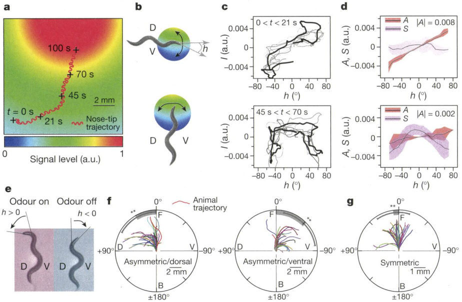


Figure 1 | Asymmetric component of the odour signal controls gradual turning. a, The trajectory of the nose tip of a nematode overlaid on a modelled exponential profile (decay length, 1 cm) of the gradient of chemoattractants (pseudocolour) from the bacterial lawn. Plus signs denote the position of the animal at the indicated times. a.u., arbitrary units. **b,** Illustrations of an animal crawling perpendicular to (top) and along (bottom) the odour gradient. h, head-bending angle; D, dorsal ($h > 0$); V, ventral ($h < 0$). **c,** Odour signal at nose tip, $I(t)$, versus head-bending angle, $h(t)$, over the last (black) and sequentially earlier (grey) head swings for the nose-tip trajectory in **a**. **d,** Plot of the mean (black line) and s.d. (coloured bar) of the asymmetric (A) and symmetric (S) components of $I(t)$ in **c**. $|A|$ denotes the magnitude of $A(t)$. **e,** Dorsal asymmetric odour stimulation (see Methods). **f, g,** Sample trajectories of the centre of mass of the animals upon dorsal and ventral asymmetric (**f**) and symmetric (**g**) odour stimulation. Angles define the turning angles with respect to the initial orientation of the animal. Grey bar denotes the mean and error bars denote s.d. over $n = 10$ experiments. B, back; F, front. $^{**}P < 0.05$, two-sample t-test.

---

[1]FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [2]Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [3]Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia 20147, USA. [4]Allen Institute for Brain Science, Seattle, Washington 98103, USA. [5]Harvard Stem Cell Institute, Harvard University, Cambridge, Massachusetts 02138, USA. [6]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.
*These authors contributed equally to this work.

signals[4,5,9,10]. Previous work on *C. elegans* has identified around 14 pairs of interneurons and motor neurons, including the interneuron pairs AIY, AIZ and AIB that are necessary for the locomotory behaviours underlying chemotaxis[3–7] (Supplementary Table 1). The neuroanatomy of the animal shows that most of the amphid chemosensory and thermosensory neurons synapse onto one or more neurons of the first layer of the interneuron pairs AIY, AIZ and AIB[11], which are further connected to a dense network of interneurons[8]. The activity dynamics in this network must process sensory signals to produce and coordinate the different locomotory behaviours underlying chemotaxis through the downstream motor neurons. Despite the experiments in the literature involving ablation, genetics and calcium imaging, we do not know whether chemotaxis is driven by key interneurons or whether the generation of this complex behaviour is achieved by the dynamics of a more diffuse neural network.

To evoke chemotactic behaviour by directly controlling interneuron activity, we have to answer the intricately linked questions of which sets of interneurons to control, and which activity patterns to stimulate in them. To deduce the classes of activity patterns triggered in the nervous system during chemotaxis, we followed animals as they crawled towards a bacterial lawn (Fig. 1a). The undulatory head swings (from dorsal to ventral, as the animal crawls on its side) caused the angle at which the head bends relative to the locomotory direction, $h(t)$ (Fig. 1b), to oscillate between positive and negative values. Owing to this changing head-bending angle and the movement of the animal, sensory cilia at the nose tip experienced the spatial profile of the chemoattractants (Fig. 1a, b) as a temporally fluctuating odour signal, $I(h,t)$. In general, this signal can be written as a sum of two terms, $I(h,t) = S_I(h,t) + A_I(h,t)$, in which S_I is a symmetric function of h: $S(h,t) = S(-h,t)$ and A_I is an asymmetric function of h:

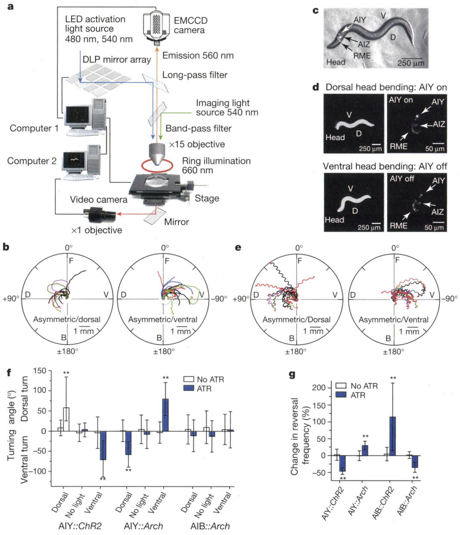


Figure 2 | Asymmetric and symmetric excitation of AIY control gradual turning and reversal frequency. **a**, Set-up for closed-loop single-neuron stimulation (see Methods). DLP, digital light processing; EMCCD, electron-multiplying charge-coupled device; LED, light-emitting diode. **b**, Sample trajectories of the centre of mass of the animals after dorsal or ventral asymmetric excitation of AIY ($n = 10$; *ttx-3::ChR2*). **c**, Fluorescence image overlaid on the bright-field image of a nematode co-expressing ChR2 and mKO in AIY, AIZ and RME neurons (*ser-2prom2::mKO* and *ser-2prom2::ChR2*). **d**, Asymmetric dorsal stimulation of the animal in **c**. On dorsal head bending (top left), ChR2 and mKO were excited in AIY using the set-up in **a** (top right, higher fluorescence in AIY) but not on ventral head bending (bottom right, decreased fluorescence in AIY). Left, ×1 dark-field images of the animal; right, ×15 fluorescent images of the neurons in the same animal. **e**, Sample trajectories of the nose tip of the animals on dorsal (left) or ventral (right) asymmetric excitation of AIY ($n = 10$; *ser-2prom2::ChR2*). **f**, Turning angle of animals after asymmetric stimulation of AIY::*ChR2* ($n = 10$), AIY::*Arch* ($n = 10$) and AIB::*Arch* ($n = 7$). ATR, all-*trans* retinal (co-factor required for functional light-gated channels *ChR2* and *Arch*). 'Dorsal' indicates asymmetric dorsal stimulation, 'ventral' indicates asymmetric ventral stimulation, 'no light' indicates unstimulated neurons, and 'no ATR' indicates control experiments. **g**, Reversal frequencies upon symmetric stimulation of AIY::*ChR2* ($n = 10$), AIY::*Arch* ($n = 19$), AIB::*ChR2* ($n = 14$) and AIB::*Arch* ($n = 11$). White and blue bars in **f** and **g** denote the mean, error bars denote 1 s.d. over n animals. $^{**}P < 0.05$, two-sample t-test).

$A(h,t) = -A(-h,t)$. When the animal moves perpendicular to the gradient direction, I is dominated by A_I (Fig. 1b–d, top). As the animal turns and tracks the gradient, the magnitude of A_I, $|A_I|$ decreases and I is dominated by S_I (Fig. 1b–d, bottom and Supplementary Fig. 1a).

To determine how these asymmetric and symmetric components of $I(t)$ control locomotory behaviour, we built a microscopy system that delivers odours on a freely crawling animal in precise temporal patterns determined by $h(t)$ (Supplementary Fig. 1b). To mimic A_I, we exposed the animal to asymmetric odour stimulation: air with chemoattractant vapours (1×10^{-3} M isoamyl alcohol) blown on the entire animal when the head was bent in one direction (for example, dorsally, $h > 0$) and odour-free air blown when the head was bent in the other direction (ventrally, $h < 0$) (Fig. 1e). The animal turned gradually in the direction in which its head was bent when the odour was delivered (Fig. 1f, Supplementary Fig. 1c, d and Supplementary Movie 1). To mimic S_I, we delivered vapours of isoamyl alcohol constantly, independent of h. The animal reduced its reversal frequency[11] and did not turn (Fig. 1g). The most parsimonious hypothesis based on these results is that asymmetric and symmetric odour components generate activity patterns in the nervous system with corresponding symmetries to control turning and reversal frequency separately during chemotaxis.

We therefore identified interneurons that triggered the different locomotory behaviours necessary for chemotaxis by directly stimulating individual neurons in a freely moving animal in asymmetric and symmetric patterns. To do so, we expressed channelrhodopsin-2 (ChR2)[12,13] or archaerhodopsin-3 (Arch)[14,15] in different neurons. Light activation of ChR2 (by 480 nm light) and Arch (540 nm) leads to neural excitation and inhibition, respectively.

Targeted illumination of specific neurons in motionless animals[16] and of body segments in freely moving animals[17,18] have been developed to excite neurons for which specific promoters are not known. Because the neurons in the nerve ring are as close as 5–10 μm to each other and their relative positions change quickly as

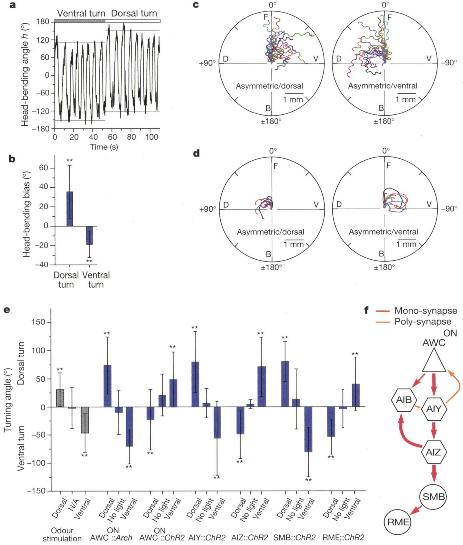


Figure 3 | Asymmetric AIY excitation modulates the head-bending angle to cause turning. a, Representative head-bending angle (horizontal lines denote means of maximum and minimum h during each full head swing) of an *AIY::ChR2* (*ttx-3::ChR2*) animal forced to first turn ventrally and then dorsally (positive angles, dorsal; negative angles, ventral). **b**, Histogram of head-bending angle bias ($\max(h) + \min(h)$) during dorsal and ventral turn caused by asymmetric AIY excitation. Error bars denote s.d. over $n = 5$. **c**, **d**, Sample tracks of animals on dorsal (left) or ventral (right) asymmetric excitation of AIZ::*ChR2* (*ser-2prom2::ChR2*) (**c**), and SMB::*ChR2* (*odr-2(18)::ChR2*) (**d**) animals. **e**, Turning angle upon asymmetric stimulation with odour ($n = 10$) and optically of AWC^{ON}::*Arch* ($n = 10$), AWC^{ON}::*ChR2* ($n = 9$), AIY::*ChR2* ($n = 10$), AIZ::*ChR2* ($n = 5$), SMB::*ChR2* ($n = 5$) and RME::*ChR2* ($n = 5$). 'Dorsal' indicates asymmetric dorsal stimulation, 'ventral' indicates asymmetric ventral stimulation and 'no light' indicates unstimulated neurons. Error bars denote s.d. over n. $^{**}P < 0.05$, two-sample t-test. **f**, The mono- and polysynapses between AWC^{ON}, AIY, AIB, AIZ, SMB and RME (thickness of arrows is proportional to synapse number). Triangle denotes sensory neuron, hexagons denote interneuron and circles denote motor neurons.

the animal moves (Supplementary Fig. 2a), we could not use these techniques. To stimulate one of many neurons optically (each with a diameter of 5–10 μm) expressing light-gated ion channels in the nerve ring of an animal that typically moves at 150 ± 50 μm s^{-1} (Supplementary Fig. 2b), our set-up tracks, identifies and specifically illuminates the neuron(s) of interest, all within 25 ms, to achieve a 5-μm spatial resolution of excitation (Fig. 2a).

Using this set-up, we first tested how stimulating the interneurons AIY and AIB affected locomotory behaviour. Both neuron pairs receive chemical synapses from the AWC sensory neurons that detect isoamyl alcohol[8] and showed calcium activity when animals were stimulated with this chemoattractant[11]. Asymmetric excitation of AIY with light in animals that expressed ChR2 only in AIY (under the promoter *ttx-3*) caused the animal to turn in the direction in which the head was bent when AIY were excited (Fig. 2b and Supplementary Movie 2). We validated our set-up by reproducing these results in animals that expressed ChR2 in AIY alone and the fluorescent protein monomeric Kusabira-Orange (mKO) in neurons AIY, AIZ and RME (under the promoter *ser-2prom2*, Supplementary Fig. 2c–f). Asymmetric stimulation of AIY in animals expressing ChR2 in AIY, AIZ and RME (*ser-2prom2*) showed the same results (Fig. 2c–f, Supplementary Fig. 2g and Supplementary Movie 3). Consistently, inhibiting the activity in AIY asymmetrically (*ttx-3::Arch*) caused the animal to turn in the opposite direction in which the head was bent when AIY were inhibited (Fig. 2f, Supplementary Fig. 3a and Supplementary Movie 4). Symmetric excitation and inhibition of AIY decreased and increased the reversal frequency respectively, but did not cause turning (Fig. 2g, Supplementary Fig. 3 and Supplementary Movie 5).

Both asymmetric and symmetric excitation or inhibition of AIB in AIB::*ChR2* and AIB::*Arch* animals (in which *ChR2* and *Arch* are expressed under the *npr-9* promoter) affected the reversal frequency of the animal but did not produce any gradual turning (Fig. 2f, g and Supplementary Fig. 4). We could thus control the two locomotory behaviours crucial for chemotaxis—gradual turns and reversal frequency—by driving different patterns of activity in AIY alone.

Turning is initiated by a larger head-bending angle in one direction[19,20]. When we forced the animal to turn by asymmetrically stimulating AIY (*ttx-3::ChR2*), the head-bending angle in the direction of the turn increased (Fig. 3a, b), indicating that asymmetric activation of AIY controlled head bending through head motor neurons to cause gradual turning. AIY neurons are most directly connected to the head muscles through the interneuron pair AIZ, which synapse onto the head motor neurons SMB and RME[8]. AIZ neurons have been proposed to have a role in gradual turns by laser ablation[4]. Ablations of SMB and RME change the head-bending angle during crawling and show loopy behaviour[7,21].

When we specifically excited AIZ::*ChR2* (*ser-2prom2::ChR2*), SMB::*ChR2* (*odr-2(18)::ChR2*) or RME::*ChR2* (*ser-2prom2::ChR2*) asymmetrically using our set-up, the animals turned (Fig. 3c–e, Supplementary Fig. 5 and Supplementary Movies 6–8). These results, in conjunction with those from asymmetric optical stimulation of the isoamyl-alcohol-sensing neuron AWC$^{\mathrm{ON}}$ (Supplementary Fig. 6 and Supplementary Movie 9), show that asymmetric stimulation of the sequence of anatomically connected neurons from AWC$^{\mathrm{ON}}$, through the interneurons AIY and AIZ to the head motor neurons SMB and RME[8] (Fig. 3f) all cause turning. These sets of neurons thus sense and respond to the component of the sensory signal that oscillates asymmetrically and in synchrony with head movement to control head bending and turning.

As different patterns of activity in AIY are sufficient to control both the frequency of reversals and turning, we tested whether controlling AIY activity alone was sufficient to coordinate reversal frequency and turning to evoke chemotactic behaviour. To do so, we measured the

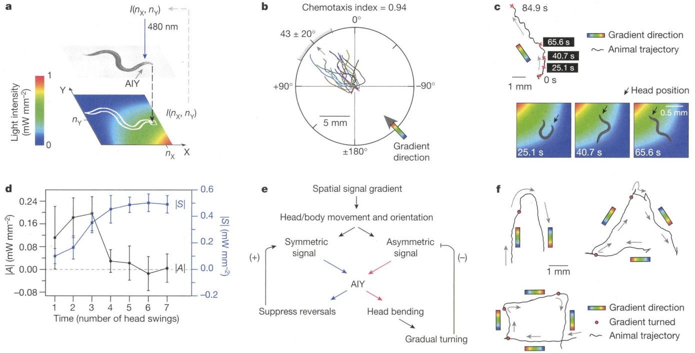


Figure 4 | Controlling AIY activity is sufficient to evoke chemotactic behaviour. **a**, Virtual light gradient algorithm (see Methods). At each time, *t*, AIY::*ChR2* animals (*ttx-3::ChR2*) are stimulated with 480 nm blue light with an intensity ($I(n_x,n_y)$) of the virtual gradient at the nose-tip position ((n_x,n_y)). **b**, Trajectories of AIY::*ChR2* animals moving in a virtual light gradient (as in **a**) with a gradient direction at 45° (black arrowheads denote the mean direction of the trajectory. Grey bar denotes s.d. ($n = 10$). **c**, Top, a sample trajectory of an animal in **b**. Bottom, snapshots of the animal making a gradual turn to reorient itself to the gradient direction (pseudocolor, same as **a**). **d**, Magnitude of A_I (black) and S_I (blue, right axis), over a head swing, as a function of the number of head swings during a gradual turn ($n = 5$, from trajectories in **b**, error bar denotes 1 s.d.). **e**, Model for chemotaxis in the virtual light gradient. **f**, Trajectories of the centre of mass of the animals when the gradient direction was suddenly rotated (at times when the animal reached the red dots) by 180°, 120° or 90°.

locomotory behaviour of animals in a fixed spatial light gradient that directly excited AIY::*ChR2* (Supplementary Fig. 7a and Supplementary Methods). The animals were unable to track the gradient direction (Supplementary Fig. 7c and Supplementary Movie 10).

The somas and processes of AIY are 150 ± 25 μm behind the nose tip of an adult animal (Supplementary Fig. 7b). As the speed of the animal is 150 ± 50 μm s^{-1}, AIY neurons follow the position of the nose tip with approximately a 1-second delay. We proposed that the dynamics of AIY excitation caused by the movement of the animal through the gradient were not in synchrony with head bending due to this delay, preventing the animals from tracking the gradient. Therefore, we designed a virtual light gradient in which the excitation light intensity on AIY depended not on the positions of AIY in space but on the position of the nose tip (Fig. 4a, Supplementary Fig. 7a and Supplementary Methods).

In this set-up, animals changed their locomotory direction using reversals and gradual turns to track the gradient direction stably (the fraction of the animals moving up the gradient and hence to the correct quadrant (defined as chemotaxis index[22]) = 0.94, Fig. 4b). As with the odour profile at the nose tip (Fig. 1d), the temporal light-intensity pattern that excited AIY could be written as a sum of an asymmetric ($A_I(h,t)$) and a symmetric ($S_I(h,t)$) component. When the locomotory direction of the animal was not along the direction of the light gradient, the magnitude of $A_I(h,t)$ over each head swing, $|A| = \sqrt{\langle A_I^2(h,t) \rangle}$, was larger, exciting AIY asymmetrically to make the animal turn. As the animal oriented itself along the direction of the gradient, the magnitude of $A_I(h,t)$ continuously diminished, suppressing turns, whereas the magnitude of $S_I(h,t)$, $|S|$, increased and suppressed reversals (Fig. 4c, d). Thus, manipulating the dynamics of activity in just the AIY interneuron pair is sufficient to evoke chemotactic behaviour. This is because the head bending and locomotion of the animal through the virtual light gradient together generate and modulate the levels of symmetric and asymmetric excitation of AIY which, in turn, control future locomotory behaviour (Fig. 4e).

This model would predict robust chemotactic behaviour in the light gradient stimulating AIY, with the symmetric and asymmetric components modulating their relative magnitudes to guide the animal stably in the correct direction. To test for robust tracking, we suddenly rotated the virtual light gradient direction by different angles and measured the response of the animal. The animals followed the gradient direction as this direction was suddenly and repeatedly rotated by 180° (Supplementary Movie 11), 120° or 90° (Fig. 4f).

Previous studies have identified neurons involved in chemotaxis by showing that defects in these neurons compromise locomotory programs and sensory modalities necessary for this behaviour. Through our approach we can identify key neurons in the neural network, the dynamics of which are sufficient to drive chemotactic behaviour and hence act as control nodes in the network. Our study leads to questions of how activity patterns stimulated in the AIY neurons during chemotaxis are disentangled by the downstream neurons to drive the different motor programs. Because many chemosensory and thermosensory neurons synapse onto AIY interneurons, it is likely that the modulation of symmetric and asymmetric activity patterns in these interneurons have a central role in the different taxis behaviours of *C. elegans*. Our techniques provide avenues to identify and generate neural activity patterns to control all of the behaviours of this nematode.

METHODS SUMMARY

Transgenic lines were constructed[2] in the *pha-1* (ref. 22) and *lite-1* (ref. 23) background and maintained using standard molecular biology techniques (see Methods). The details of the data analysis to quantify chemotactic behaviour, reversal frequency and the set-up for odour stimulation, single neuron stimulation and virtual light gradients are explained in Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 17 May; accepted 23 July 2012.
Published online 23 September 2012.

1. Brenner, S. The genetics of behaviour. *Br. Med. Bull.* **29,** 269–271 (1973).
2. Brenner, S. The genetics of *Caenorhabditis elegans. Genetics* **77,** 71–94 (1974).
3. Tsalik, E. L. & Hobert, O. Functional mapping of neurons that control locomotory behavior in *Caenorhabditis elegans. J. Neurobiol.* **56,** 178–197 (2003).
4. Iino, Y. & Yoshida, K. Parallel use of two behavioral mechanisms for chemotaxis in *Caenorhabditis elegans. J. Neurosci.* **29,** 5370–5380 (2009).
5. Pierce-Shimomura, J. T., Morse, T. M. & Lockery, S. R. The fundamental role of pirouettes in *Caenorhabditis elegans* chemotaxis. *J. Neurosci.* **19,** 9557–9569 (1999).
6. Wakabayashi, T., Kitagawa, I. & Shingai, R. Neurons regulating the duration of forward locomotion in *Caenorhabditis elegans. Neurosci. Res.* **50,** 103–111 (2004).
7. Gray, J. M., Hill, J. J. & Bargmann, C. I. A circuit for navigation in *Caenorhabditis elegans. Proc. Natl Acad. Sci. USA* **102,** 3184–3191 (2005).
8. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans. Phil. Trans. R. Soc. Lond. B* **314,** 1–340 (1986).
9. Ward, S. Chemotaxis by the nematode *Caenorhabditis elegans*: identification of attractants and analysis of the response by use of mutants. *Proc. Natl Acad. Sci. USA* **70,** 817–821 (1973).
10. Izquierdo, E. J. & Lockery, S. R. Evolution and analysis of minimal neural circuits for klinotaxis in *Caenorhabditis elegans. J. Neurosci.* **30,** 12908–12917 (2010).
11. Chalasani, S. H. *et al.* Dissecting a circuit for olfactory behaviour in *Caenorhabditis elegans. Nature* **450,** 63–70 (2007).
12. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8,** 1263–1268 (2005).
13. Nagel, G. *et al.* Light activation of channelrhodopsin-2 in excitable cells of *Caenorhabditis elegans* triggers rapid behavioral responses. *Curr. Biol.* **15,** 2279–2284 (2005).
14. Chow, B. Y. *et al.* High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* **463,** 98–102 (2010).
15. Okazaki, A., Sudo, Y. & Takagi, S. Optical silencing of *C. elegans* cells with arch proton pump. *PLoS ONE* **7,** e35370 (2012).
16. Guo, Z. V., Hart, A. C. & Ramanathan, S. Optical interrogation of neural circuits in *Caenorhabditis elegans. Nature Methods* **6,** 891–896 (2009).
17. Leifer, A. M., Fang-Yen, C., Gershow, M., Alkema, M. J. & Samuel, A. D. T. Optogenetic manipulation of neural activity in freely moving *Caenorhabditis elegans. Nature Methods* **8,** 147–152 (2011).
18. Stirman, J. N. *et al.* Real-time multimodal optical control of neurons and muscles in freely behaving *Caenorhabditis elegans. Nature Methods* **8,** 153–158 (2011).
19. Lockery, S. R. The computational worm: spatial orientation and its neuronal basis in *C. elegans. Curr. Opin. Neurobiol.* **21,** 782–790 (2011).
20. Kim, D., Park, S., Mahadevan, L. & Shin, J. H. The shallow turn of a worm. *J. Exp. Biol.* **214,** 1554–1559 (2011).
21. McIntire, S. L., Jorgensen, E., Kaplan, J. & Horvitz, H. R. The GABAergic nervous system of *Caenorhabditis elegans. Nature* **364,** 337–341 (1993).
22. Granato, M., Schnabel, H. & Schnabel, R. *pha-1*, a selectable marker for gene transfer in *C. elegans. Nucleic Acids Res.* **22,** 1762–1763 (1994).
23. Edwards, S. L. *et al.* A novel molecular solution for ultraviolet light detection in *Caenorhabditis elegans. PLoS Biol.* **6,** e198 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Dowling, S. Lockery, J. Lichtman, K. McCormick, A. Murray, E. O'Shea, A. Schier, B. Stern and members of the Ramanathan laboratory for discussions and comments, the Human Frontier Science Program (HFSP) Postdoctoral Fellowship (A.K.), National Science Foundation (NSF) Graduate Fellowship (C.-H.S.), NSF Career Award, Pew Scholar, Klingenstein Fellowship Award and the National Institutes of Health (NIH) Pioneer Awards (S.R.) for support.

Author Contributions A.K., C.-H.S., Z.V.G. and S.R. designed the experiments. A.K., C.-H.S. and Z.V.G. performed the experiments. A.K., C.-H.S. and S.R. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.K. (akocabas@cgr.harvard.edu) or S.R. (sharad@post.harvard.edu).

doi:10.1038/nature11415

Restoration of auditory evoked responses by human ES-cell-derived otic progenitors

Wei Chen[1,2]*, Nopporn Jongkamonwiwat[1,2,3]*, Leila Abbas[1,2], Sarah Jacob Eshtan[1,2], Stuart L. Johnson[2], Stephanie Kuhn[2], Marta Milo[2], Johanna K. Thurlow[1,2], Peter W. Andrews[1,2], Walter Marcotti[2], Harry D. Moore[1,2] & Marcelo N. Rivolta[1,2]

Deafness is a condition with a high prevalence worldwide, produced primarily by the loss of the sensory hair cells and their associated spiral ganglion neurons (SGNs). Of all the forms of deafness, auditory neuropathy is of particular concern. This condition, defined primarily by damage to the SGNs with relative preservation of the hair cells[1], is responsible for a substantial proportion of patients with hearing impairment[2]. Although the loss of hair cells can be circumvented partially by a cochlear implant, no routine treatment is available for sensory neuron loss, as poor innervation limits the prospective performance of an implant[3]. Using stem cells to recover the damaged sensory circuitry is a potential therapeutic strategy. Here we present a protocol to induce differentiation from human embryonic stem cells (hESCs) using signals involved in the initial specification of the otic placode. We obtained two types of otic progenitors able to differentiate *in vitro* into hair-cell-like cells and auditory neurons that display expected electrophysiological properties. Moreover, when transplanted into an auditory neuropathy model, otic neuroprogenitors engraft, differentiate and significantly improve auditory-evoked response thresholds. These results should stimulate further research into the development of a cell-based therapy for deafness.

Hair-cell-like phenotypes and sensory neurons, with different degrees of functional maturation, have been obtained from mouse stem populations[4–10]. After transplantation, some cell types have shown engraftment but none have shown evidence of functional recovery[10–15]. Although useful for research purposes, these products are unsuitable for a therapeutic application and appropriate cell types of human origin have remained elusive so far. Neuroprogenitors isolated from mature human cochleae display limited proliferative and differentiating potential[16], hESC-derived neural crest cells may differentiate into sensory neurons by exposure to bone morphogenetic protein (BMP) but lack true otic characteristics[17,18]. Recently, we isolated a population of bipotent stem cells from the human fetal cochlea (human fetal auditory stem cells, hFASCs), with the ability to produce hair-cell-like cells and neurons[19]. However, although hFASCs can be expanded *in vitro* for approximately 25 population doublings, they eventually undergo replicative senescence. Hence, there is a need for a reliable, renewable source of human otic progenitors, with the ability to produce both cell types for sensory replacement.

Fibroblast growth factor (FGF) signalling is necessary and sufficient for the induction *in vivo* of the otic placode, the primordium of the hearing organ[20,21]. As the ligands involved in placode signalling in the mouse have been identified as FGF3 and FGF10 (refs 22, 23), we proposed that exposure to these factors would trigger otic differentiation of hESCs. Initial experiments with embryoid bodies confirmed FGF3 and FGF10 induction of otic features (Supplementary Fig. 1a). We therefore focused on developing a method devoid of this initial cell-aggregation step, which is prone to high variability. Undifferentiated colonies of hESCs were dissociated for plating as a monolayer on laminin-coated flasks (see Supplementary Methods). Under these conditions, treatment with FGF3 and FGF10 induced the placodal markers *PAX8* and *PAX2,* either in the presence of knockout serum replacement or under defined conditions using Dulbecco's Modified Eagle's Medium (DMEM) with Ham's F12 and N2/B27 (DFNB) medium (Supplementary Methods and Supplementary Figs 1b and 2). Global analyses of gene expression was performed using Affymetrix GeneChip arrays and, after normalization (see Supplementary Methods), samples were mined in two different ways. We first used the Gene Set Enrichment Analysis (GSEA) tool[24] to look for genes that were enriched in the entire list of probe sets, without establishing a priori cut off of differential expression (Supplementary Tables 1 and 2). This analysis showed that a set of otic markers was significantly enriched in the FGF-treated samples when compared with the undifferentiated hESCs (normalized enriched score (NES), 0.568; family-wise error rate (FWER) $P = 0.046$) or cells grown in DFNB (NES, 0.707; FWER $P = 0.019$) (Supplementary Table 1). A second type of analysis assessed genes differentially expressed using predefined criteria for fold-change cut off and statistical significance (see Supplementary Methods). A total of 1,424 genes (represented by 2,124 probe sets) was differentially upregulated in the FGF samples when compared to undifferentiated hESCs, whereas 423 genes (505 probe sets) were unregulated in the FGF-treated versus the DFNB controls (Supplementary Tables 3 and 4). Conversely, 2,368 genes (3,231 probe sets) were downregulated in the FGF samples versus hESCs, and 482 genes (607 probe sets) were downregulated versus DFNB (Supplementary Tables 5 and 6). In a gene ontology analysis, the gene ontology terms 'sensory organ development' (Expression Analysis Systematic Explorer (EASE) P value score in FGF versus hESC, $P = 3.92 \times 10^{-15}$; FGF versus DFNB, $P = 0.022$); 'ear development' (FGF versus hESC, $P = 4.47 \times 10^{-8}$; FGF versus DFNB, $P = 0.014$) and 'ear morphogenesis' (FGF versus hESC, $P = 3.08 \mathrm{x} 10^{-6}$; FGF versus DFNB, $P = 0.0497$) were highly enriched in the FGF-treated cells in both comparisons, and 'mechanoreceptor differentiation' and 'auditory receptor differentiation' were highly enriched in FGF versus hESC (see Supplementary Tables 7–10). Both bioinformatics analyses therefore suggested that the FGF treatment was generating a global change of transcription compatible with the induction of otic progenitors.

We also used immunostaining to examine the co-expression of PAX8 and SOX2, to define the otic progenitors at a cellular level. Otic progenitors grew as colonies after the inductive phase. Initial immunolabelling showed a relatively large proportion of double-positive cells in the FGF-treated condition (~78%), in contrast to the relatively moderate upregulation of otic transcripts detected with the arrays. However, a subset of cells expressed very high levels of PAX8 and SOX2, and these were assessed with an automated microscopy platform (InCell Analyzer 1000) that enabled quantification of the number of positive cells and their relative intensity (Fig. 1 and Supplementary Fig. 3). When a stringent threshold was selected

---

[1]Centre for Stem Cell Biology, University of Sheffield, Sheffield S10 2TN, UK. [2]Department of Biomedical Sciences, University of Sheffield, Sheffield S10 2TN, UK. [3]Faculty of Health Sciences, Srinakharinwirot University, Ongkharak, Nakhonnayok 26120, Thailand.

*These authors contributed equally to this work.

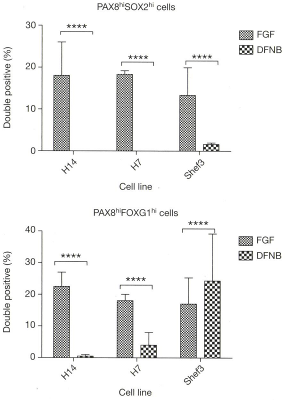


Figure 1 | FGF3 and FGF10 generate otic progenitors. Bar charts showing the percentage of PAX8^{hi}SOX2^{hi}-expressing (top panel) and PAX8^{hi}FOXG1^{hi}-expressing (bottom panel) cells at the FGF 75th percentile threshold, obtained from the hESC lines H14, H7 and Shef3 ($n = 3$). Error bars, mean + s.e.m. $^{****}P < 0.0001$.

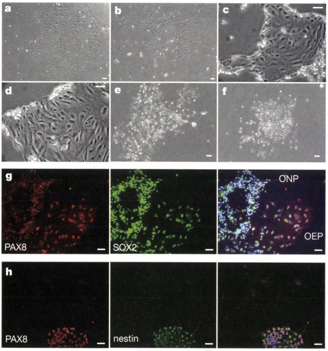


Figure 2 | Otic epithelial progenitors and otic neuro progenitors. **a**, **b**, Morphology of an OEP colony. Scale bars, 100 μm. **c**, **d**, The partial lifting of OEPs when treated with a short, mild trypsin incubation. Scale bars, 50 μm **e**, **f**, Typical morphology of ONPs, showing cytoplasmic projections. Scale bars, 50 μm. **g**, Side-by-side ONP and OEP colonies, double-labelled for PAX8 and SOX2. Scale bars, 50 μm. **h**, ONP colony labelled for PAX8 and nestin. Scale bars, 50 μm.

(75th intensity percentile per cell line and antibody; see Supplementary Methods) 18.3% ± 0.8 (± s.e.m.) of the cells expressed high levels of PAX8 and SOX2 (PAX8^{hi}SOX2^{hi}) after FGF treatment (against 0% obtained without the growth factors, $P < 0.001$), whereas 18% ± 2 cells were PAX8^{hi}FOXG1^{hi} (compared to 4% ± 4 cells for the control, $P < 0.001$). PAX8^{hi}SOX2^{hi}FOXG1^{hi} cells also expressed the otic markers PAX2, nestin, SIX1 and GATA3 (Fig. 2h and Supplementary Figs 4 and 5a). It is likely that this subset of PAX8^{hi}SOX2^{hi}FOXG1^{hi}-expressing cells represents the otic progenitors. The reproducibility of the protocol was tested across the hESC lines H7, H14 and Shef3, which all gave comparable results (see Fig. 1 and Supplementary Fig. 3). FGF3 and FGF10 induced two morphologically distinct types of otic colonies (Fig. 2a–h). One cell population showed a flat phenotype, with large cytoplasm and formed epithelioid islands (Fig. 2a–d), whereas the second was small, with denser chromatin, and presented cytoplasmic projections (Fig. 2e, f). Given their morphological appearance, we have operationally named them otic epithelial progenitors (OEPs) and otic neural progenitors (ONPs), respectively. The relative proportion of these progenitors was dependent on the cell line, plating density and the degree of cell separation (single cells versus cell clusters) (Supplementary Figs 5 and 6 and Supplementary Methods). Progenitor colonies were purified using sequential dissociation (see Supplementary Methods), yielding moderately homogenous cultures of the desired cell colony type, and were expanded in OSCFM (otic stem cell full media; see Supplementary Methods).

The differentiation potential of OEPs and ONPs was tested in 'neuralizing' and 'hair-cell' culture conditions developed previously using hFASCs[19] (see Supplementary Methods). OEPs produced hair-cell-like cells as defined by the simultaneous expression of ATOH1 and BRN3C, or BRN3C and MYO7A (~45%) (Supplementary Fig. 7). A small subset differentiated a rudimentary apical bundle, expressing espin (Supplementary Fig 8). These hair-cell-like cells also expressed an outward K$^+$ current, the inward rectifier K$^+$ current I_{K1} and an inward Ca^{2+} current (I_{Ca}) (Supplementary Fig. 9). Under neuralizing conditions, they produced a small proportion (~9%) of sensory neurons (Supplementary Fig. 7). Conversely, ONPs were committed to produce neurons. Under neuralizing conditions, almost all cells developed a bipolar morphology and were positive for BRN3A and β-tubulin III, as well as for β-tubulin III and NF200. They also expressed *NEUROD1*, *ISL1* and *NTRK2*, a delayed-rectifier K$^+$ current (I_K), an Na$^+$ current (I_{Na}), and elicited single action potentials (Supplementary Fig. 9). No hair-cell differentiation was obtained from ONPs under neuralizing or hair-cell culture conditions. Detailed results are given in Supplementary Information.

The properties of ONPs *in vivo* were studied by transplanting them into ouabain-treated gerbils, a model of neuropathic deafness[25]. Application of ouabain directly to the round window selectively damages the type I SGNs, preserving the hair cells and the organ of Corti[26] (Supplementary Fig. 10). After ouabain application, only a small number of SGNs survived (6.4%; see Supplementary Table 11). Most of the surviving cells (~87%) were peripherin-positive type II neurons, therefore less than 1% of the original population of type I neurons remained (Supplementary Table 11 and Supplementary Fig. 13). Staining for myosin VIIa and the presence of distortion product otoacoustic emissions (DPOAEs) confirmed that the organ of Corti had not been damaged (Supplementary Figs 10 and 11). DPOAEs are sounds produced as a consequence of electromechanical feedback from the outer hair cells and can be used to check their physiological integrity.

ONPs derived from Shef1 hESCs constitutively expressing either enhanced green fluorescent protein (eGFP) or tomato fluorescent protein were expanded in OSCFM, dissociated with trypsin and delivered directly into the modiolus, approaching the cochlea through the round window. One set of animals was transplanted 3 to 5 days after ouabain application ($n = 13$), and another was transplanted 2 weeks after the ototoxic drug ($n = 5$). As no functional or histological differences were encountered between the two groups ($P > 0.05$; Supplementary Fig. 12), they were analysed together. Two to three weeks after transplantation, five out of six animals had surviving, transplanted cells grafted in the modiolus, forming an ectopic spiral ganglion (Fig. 3a, b). Cells in the marginal sides of the ectopic ganglion had undergone differentiation as judged by β-tubulin III staining (Fig. 3b) and displayed neural projections, targeting the organ of Corti (Fig. 3c, d). Animals were then monitored for 10 weeks post tranplantation. Histological analysis after 10 weeks post transplantation showed that the ectopic ganglion was still present and that cells had also migrated into the Rosenthal's canal (Fig. 3e). Transplanted cells expressed the 3A10 neurofilament-associated antigen and NKAα3 (ATP1A3), a marker of type I neurons and afferent fibres in the inner ear[27] (Supplementary Fig. 14). Notably, projections from the transplanted cells that reached the organ of Corti were targeting the hair cells, and fibres positive for NKAα3 and GluA2 were found next to the basal pole of the inner hair cells, suggesting the presence of synaptic connections (Fig. 3g). Moreover, fibres form the transplanted cells were visualized leaving the modiolus towards the brainstem (Fig. 3f). In the cochlear nucleus of three gerbils we found red fluorescent protein (RFP)-positive fibres also stained for synaptophysin, suggesting synaptic connections with the central auditory path (Fig. 3h, i). Transplanted ONPs contributed significantly to restore neuronal density (Fig. 3j; $P < 0.01$). Although 112.5 ± 11.9 TUJ1$^+$ cells per mm^2 (± s.e.m.) were present in the ouabain-treated, untransplanted ears, 546.4 ± 30.6 TUJ1$^+$ cells per mm^{-2} were found after transplantation. From these, 94.9 ± 0.3% were also GFP- (or tomato)-positive, confirming their exogenous nature (Supplementary Table 12). The number of projections detected in the brainstem was considerably lower than the number of transplanted cell bodies identified in the ganglion. Although this could be explained by the limited sorting of fluorescent protein into the long afferent fibres, the pathfinding of the central innervations requires further future exploration. No tumours were detected in any of the transplanted animals at any stage throughout the experiment.

Functional performance was determined by measuring auditory-evoked response (ABR) thresholds. These were established based on the wave ii–wave iii (P2–N3 waveform) amplitude[28]. These waves are generated by the cochlear nucleus and the superior olivary complex cells, and reflect neural connections with the central auditory pathway[29]. After ouabain application, auditory function was severely impaired, with thresholds rising from 20 dB sound pressure level (SPL) to almost 80 dB SPL, the maximum intensity tested. Frequency discrimination was also abolished. The amplitudes of wave ii–wave iii complexes were almost negligible at any of the frequencies explored at the maximum intensity of 80 dB SPL (Fig. 4d). ABRs were recorded at 1- to 2-week intervals. Control animals ($n = 8$) showed no sign of functional recovery throughout the experiment, with a mean auditory threshold after 10 weeks of 75.14 ± 2.3 dB (± s.e.m.); similar to the 76.37 ± 1.8 dB obtained after ouabain treatment. However, in the transplanted animals ($n = 18$), there was a detectable improvement in the ABR thresholds (Fig. 4a, b) starting approximately 4 weeks post transplantation, with the

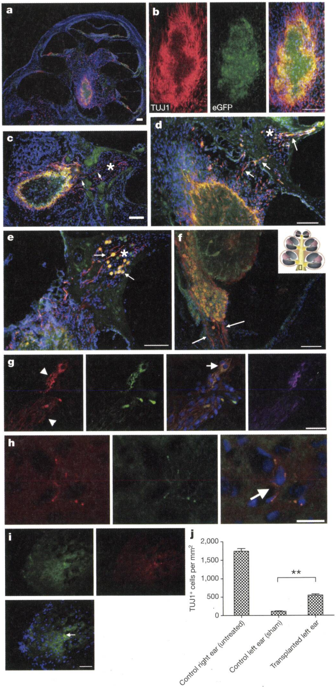


Figure 3 | Transplantation of otic progenitors restores a population of spiral ganglion neurons. **a**, Mid-modiolar section of a transplanted cochlea showing the location of the newly formed, ectopic ganglion. **b**, Detail of the ganglion showing neuronal differentiation by TUJ1 staining (left panel), with eGFP (middle panel) and an overlaid image (right panel). **c**, **d**, Neural fibres project from the ganglion towards the organ of Corti (arrows), passing through the Rosenthal's canal (asterisk). **e**, New neuronal bodies (arrows) are also found in the Rosenthal's canal (asterisk). **f**, Ectopic ganglion at the base of the modiolus, projecting TUJ1$^+$ fibres centrally towards the internal auditory meatus. Inset, schematic of the cochlea showing the location of the ectopic ganglion. **g**, RFP$^+$ fibres (left panel; arrowheads) approaching the inner hair cells and expressing GluA2 (middle-left panel), primarily concentrated in postsynaptic densities (PSDs) around the basal pole of inner hair cells (IHCs) (middle-right panel; arrow). Dotted lines show the positions of the IHCs. Fibres (including PSDs) were also positive for NKAα3 (also known as ATP1A3; right panel), a marker of afferent terminals. Nine out of ten animals analysed had fibres contacting the IHC, whereas the three animals labelled for GluA2 were positive for this marker. **h**, **i**, RFP$^+$ fibres in the cochlear nucleus, expressing synaptophysin (arrows). The fibre branches and surrounds the cell (**h**), with morphology highly reminiscent of the maturing endbulb of Held. Left, RFP; middle, synapthophysin; right, overlaid image. **j**, SGN density 10 weeks after transplantation. Conditions compared are cochleae treated with ouabain and sham-operated (control left ear (sham)) versus those treated with ouabain and transplanted with ONPs (transplanted left ear). Density was significantly increased ($P < 0.01$) from 112.5 ± 11.9 (control left ear (sham); $n = 3$; mean + s.e.m.) to 546.4 ± 30.6 (transplanted left ear; $n = 8$). As a reference, the density of control, right ear untreated cochleae, was 1,743 ± 71.5 TUJ1$^+$ cells per mm^2. $^{**}P < 0.01$. Scale bars for **a–f**, 100 μm; for **g–i**, 50 μm.

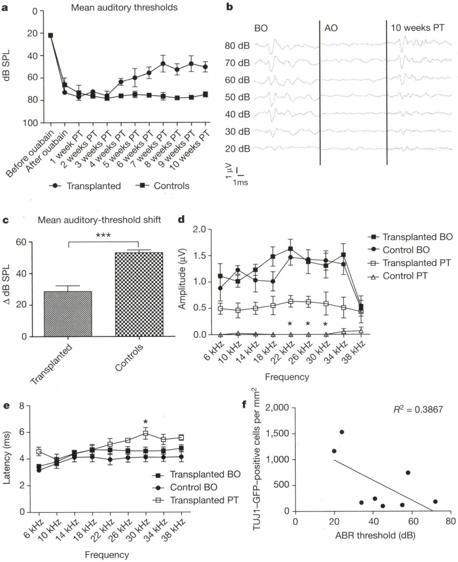


Figure 4 | Transplanted cells provide a recovery of ABR thresholds. **a**, Evolution of the mean ABR thresholds (click) obtained in the transplanted animals ($n = 18$; mean ± s.e.m.) compared to the controls ($n = 8$). **b**, Trace ABR showing the abolition of waves after ouabain (AO) treatment and the restoration of the complexes 10 weeks post transplanation (PT). **c**, Graph showing the mean auditory-threshold shift reduction obtained by the transplantation (transplanted 28.6 ± 3.6 dB; $n = 18$ versus 53 ± 1.7 dB; $n = 8$ in the control, $P = 0.0002$; mean + s.e.m). **d**, Comparison of the wave ii–wave iii amplitudes obtained by tone ABRs. A general trend of enhanced amplitudes was obtained across all frequencies tested, being significantly ($^*P < 0.05$) different from the untransplanted controls at 22, 26 and 30 kHz. Amplitudes before ouabain were equivalent between the transplanted ($n = 6$) and untransplanted animals ($n = 5$; mean ± s.e.m.). **e**, Latencies of wave ii–wave iii complexes were, in general, comparable before ouabain and after transplantation. Only at 30 kHz was a significant delay was observed (BO: 4.58 ± 0.2 ms, $n = 6$; PT: 5.9 ± 0.4 ms, $n = 5$; $P < 0.05$; mean ± s.e.m.). **f**, A significant correlation was observed between the mean density of TUJ1–GFP-positive cells and the ABR thresholds ($n = 8$; $P < 0.05$).
$^*P < 0.05$; $^{***}P < 0.001$.

mean auditory threshold lowered (improved) to 50.4 ± 4.5 dB by 10 weeks post transplantation. Furthermore, the mean auditory threshold shift, calculated as the difference between the threshold at 10 weeks post transplantation versus the one before ouabain treatment, was of 53 ± 1.7 dB in the control animals, compared to 28.6 ± 3.6 dB in the transplanted cohort ($P = 0.0002$; Fig. 4c). This represents an overall functional improvement of approximately 46%. The range of recovery went from modest to almost complete (see Supplementary Fig. 15), which is remarkable considering the technical challenges involved in the procedure. Tonotopical processing was also partially restored (Fig. 4d). A trend in the increment of wave ii–wave iii amplitudes was detected at each frequency explored, with amplitudes being significantly different at 22, 26 and 30 kHz, when compared to the untransplanted animals ($P < 0.05$). When compared to the amplitudes before the ouabain application, the improvement was approximately 43%. Latencies were mostly similar to the ones before ouabain (Fig. 4e). The only significant difference was detected at 30 kHz (before ouabain, 4.58 ± 0.2 ms, $n = 6$; post transplantation, 5.9 ± 0.4 ms, $n = 5$; $P < 0.05$), suggesting that some maturation was still taking place at this stage. Finally, there was a significant correlation between the increment of neural density by transplanted cells and the lowering of the ABR threshold ($R^2 = 0.3867$, $P < 0.05$; Fig. 4f).

Our developmentally informed protocol produced hESC-derived auditory hair cells and neurons that closely resembled phenotypes obtained from hFASCs, providing validation of their cochlear characteristics. This was supported further by the restoration of ABR thresholds on transplantation of otic progenitors into a deaf adult mammal. The ability to reinstate auditory neuron functionality paves the way for a future cell-based treatment for auditory neuropathies. It may also, in combination with a cochlear implant, offer a therapeutic solution to a wider range of patients that currently remain without viable treatment.

METHODS SUMMARY

hESCs lines used (H7, H14, Shef1, Shef3, Shef1–eGFP and Shef1–tomato), with a normal karyotype, were maintained on mouse embryonic fibroblast feeders (MEFs) under standard conditions. Although experiments with embryoid bodies and initial monolayer experiments were carried out in the presence of KOSR, we later adopted a chemically defined medium. This serum-free, chemically defined basal culture media included a 1:1 mixture of Dulbeco's Modified Eagle's Medium (DMEM):Ham's F12 and N2/B27 supplements. In most experiments, FGF3 and FGF10 were used at 50 $\mathrm{ng\,ml^{-1}}$. Laminin (R&D Systems) was used at 5 $\mathrm{\mu g\,cm^{-2}}$. Antibodies, polymerase chain reaction (PCR) primers and microarray analysis are detailed in the Supplementary Methods. To induce hair-cell differentiation, progenitors were transferred to gelatin-coated dishes and cultured with DFNB supplemented with all-*trans* retinoic acid (10^{-6} M; Sigma) and epidermal growth factor (EGF) (20 $\mathrm{ng\,ml^{-1}}$) for 2 to 4 weeks. To induce neuronal differentiation, cells dissociated with trypsin were plated on gelatin-coated dishes and incubated in DFNB with basic FGF (20 $\mathrm{ng\,ml^{-1}}$) and Sonic hedgehog (Shh-C24II, 500 $\mathrm{ng\,ml^{-1}}$; R&D Systems). On the third day, culture was supplemented with neurotrophin 3 (NTF3, 10 $\mathrm{ng\,ml^{-1}}$; Petropech) and brain-derived neurotrophic factor (BDNF, 10 $\mathrm{ng\,ml^{-1}}$; Petropech). Shh-C24II was removed on the fourth or fifth day, whereas the neurotrophins remained for the length of the incubation, normally between 7 and 14 days. Conditions for electrophysiological recordings are detailed in the Supplementary Methods. The auditory neuropathy model was generated by applying 1 mM ouabain directly into the round-window niche of adult gerbils. Either 3 days or 2 weeks later, hONPs expressing eGFP or tomato fluorescent protein, were injected into the modiolus. Functional recovery was monitored weekly by measuring ABRs and DPOAEs, for up to 10 weeks. Cochleae were taken, fixed and processed for analysis. Details of the hearing test and histological preparation are provided in the Supplementary Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 31 July 2010; accepted 16 July 2012.

Published online 12 September 2012.

1. Vlastarakos, P. V., Nikolopoulos, T. P., Tavoulari, E., Papacharalambous, G. & Korres, S. Auditory neuropathy: endocochlear lesion or temporal processing impairment? Implications for diagnosis and management. *Int. J. Pediatr. Otorhinolaryngol.* **72,** 1135–1150 (2008).
2. Uus, K. & Bamford, J. Effectiveness of population-based newborn hearing screening in England: ages of interventions and profile of cases. *Pediatrics* **117,** e887–e893 (2006).
3. Bradley, J., Beale, T., Graham, J. & Bell, M. Variable long-term outcomes from cochlear implantation in children with hypoplastic auditory nerves. *Cochlear Implants Int.* **9,** 34–60 (2008).
4. Li, H., Liu, H. & Heller, S. Pluripotent stem cells from the adult mouse inner ear. *Nature Med.* **9,** 1293–1299 (2003).
5. Li, H., Roblin, G., Liu, H. & Heller, S. Generation of hair cells by stepwise differentiation of embryonic stem cells. *Proc. Natl Acad. Sci. USA* **100,** 13495–13500 (2003).
6. Oshima, K. *et al.* Mechanosensitive hair cell-like cells from embryonic and induced pluripotent stem cells. *Cell* **141,** 704–716 (2010).
7. Jeon, S. J., Oshima, K., Heller, S. & Edge, A. S. Bone marrow mesenchymal stem cells are progenitors *in vitro* for inner ear hair cells. *Mol. Cell. Neurosci.* **34,** 59–68 (2007).
8. Kondo, T., Johnson, S. A., Yoder, M. C., Romand, R. & Hashino, E. Sonic hedgehog and retinoic acid synergistically promote sensory fate specification from bone marrow-derived pluripotent stem cells. *Proc. Natl Acad. Sci. USA* **102,** 4789–4794 (2005).
9. Coleman, B., Fallon, J. B., Pettingill, L. N., de Silva, M. G. & Shepherd, R. K. Auditory hair cell explant co-cultures promote the differentiation of stem cells into bipolar neurons. *Exp. Cell Res.* **313,** 232–243 (2007).
10. Reyes, J. H. *et al.* Glutamatergic neuronal differentiation of mouse embryonic stem cells after transient expression of neurogenin 1 and treatment with BDNF and GDNF: *in vitro* and *in vivo* studies. *J. Neurosci.* **28,** 12622–12631 (2008).
11. Corrales, C. E. *et al.* Engraftment and differentiation of embryonic stem cell-derived neural progenitor cells in the cochlear nerve trunk: growth of processes into the organ of Corti. *J. Neurobiol.* **66,** 1489–1500 (2006).
12. Hildebrand, M. S. *et al.* Survival of partially differentiated mouse embryonic stem cells in the scala media of the guinea pig cochlea. *J. Assoc. Res. Otolaryngol.* **6,** 341–354 (2005).
13. Sekiya, T. *et al.* Transplantation of conditionally immortal auditory neuroblasts to the auditory nerve. *Eur. J. Neurosci.* **25,** 2307–2318 (2007).
14. Lang, H. *et al.* Transplantation of mouse embryonic stem cells into the cochlea of an auditory-neuropathy animal model: effects of timing after injury. *J. Assoc. Res. Otolaryngol.* **9,** 225–240 (2008).
15. Hu, Z., Ulfendahl, M. & Olivius, N. P. Central migration of neuronal tissue and embryonic stem cells following transplantation along the adult auditory nerve. *Brain Res.* **1026,** 68–73 (2004).
16. Rask-Andersen, H. *et al.* Regeneration of human auditory nerve. In vitro/in video demonstration of neural progenitor cells in adult human and guinea pig spiral ganglion. *Hear. Res.* **203,** 180–191 (2005).
17. Shi, F., Corrales, C. E., Liberman, M. C. & Edge, A. S. BMP4 induction of sensory neurons from human embryonic stem cells and reinnervation of sensory epithelium. *Eur. J. Neurosci.* **26,** 3016–3023 (2007).
18. Lee, G. *et al.* Isolation and directed differentiation of neural crest stem cells derived from human embryonic stem cells. *Nature Biotechnol.* **25,** 1468–1475 (2007).
19. Chen, W. *et al.* Human fetal auditory stem cells can be expanded *in vitro* and differentiate into functional auditory neurons and hair cell-like cells. *Stem Cells* **27,** 1196–1204 (2009).
20. Martin, K. & Groves, A. K. Competence of cranial ectoderm to respond to Fgf signaling suggests a two-step model of otic placode induction. *Development* **133,** 877–887 (2006).
21. Freter, S., Muta, Y., Mak, S. S., Rinkwitz, S. & Ladher, R. K. Progressive restriction of otic fate: the role of FGF and Wnt in resolving inner ear potential. *Development* **135,** 3415–3424 (2008).
22. Alvarez, Y. *et al.* Requirements for FGF3 and FGF10 during inner ear formation. *Development* **130,** 6329–6338 (2003).
23. Wright, T. J. & Mansour, S. L. Fgf3 and Fgf10 are required for mouse otic placode induction. *Development* **130,** 3379–3390 (2003).
24. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
25. Schmiedt, R. A., Okamura, H. O., Lang, H. & Schulte, B. A. Ouabain application to the round window of the gerbil cochlea: a model of auditory neuropathy and apoptosis. *J. Assoc. Res. Otolaryngol.* **3,** 223–233 (2002).
26. Lang, H., Schulte, B. A. & Schmiedt, R. A. Ouabain induces apoptotic cell death in type I spiral ganglion neurons, but not type II neurons. *J. Assoc. Res. Otolaryngol.* **6,** 63–74 (2005).
27. McLean, W. J., Smith, K. A., Glowatzki, E. & Pyott, S. J. Distribution of the Na,K-ATPase alpha subunit in the rat spiral ganglion and organ of corti. *J. Assoc. Res. Otolaryngol.* **10,** 37–49 (2009).
28. Burkard, R., Boettcher, F., Voigt, H. & Mills, J. Comments on "Stimulus dependencies of the gerbil brain-stem auditory-evoked response (BAER). I: Effects of click level, rate and polarity" [J. Acoust. Soc. Am. 85, 2514–2525 (1989)]. *J. Acoust. Soc. Am.* **94,** 2441–2442 (1993).
29. Boettcher, F. A., Mills, J. H. & Norton, B. L. Age-related changes in auditory evoked potentials of gerbils. I. Response amplitudes. *Hear. Res.* **71,** 137–145 (1993).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported primarily by grants from Action on Hearing Loss (RNID) to M.N.R. Other support included Deafness Research UK (M.N.R and W.M.), the Wellcome Trust (088719, W.M.), Medical Research Council (P.W.A, H.D.M. and M.N.R) and ESTOOLS (P.W.A). S.L.J. was supported by a Wellcome Trust VIP award and the RNID. W.M. and S.L.J. are Royal Society university research fellows. Confocal images were taken at the Light Microscopy Facility of the Department of Biomedical Sciences, University of Sheffield. We are grateful for the advice of M. Mulheran and I. Russell on the tests of auditory function, provided at the earlier stages of this project, and to the assistance of P. Gokhale on the use of the InCell Analyzer. M.N.R. acknowledges the support and encouragement of his late parents, Noemí Luján-Ceballos and Juan Carlos Rivolta.

Author Contributions W.C., N.J., L.A., S.J.E and J.K.T. collected and/or assembled, analysed and interpreted data. S.L.J., S.K. and W.M. collected and/or assembled, analysed and intepreted electrophysiology data. M.M. carried out biocomputational analysis of gene-array data. P.W.A. and H.D.M. provided study material and administrative support. M.N.R. was responsible for conceiving and designing the study, for obtaining financial support, collecting and/or assembling, analysing and interpreting data, for manuscript writing, and for final approval of the manuscript.

Author Information Microarray datasets have been deposited at the NCBI Gene Expression Omnibus and they can be retrieved with accession number GSE36754. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.N.R. (m.n.rivolta@sheffield.ac.uk).

LETTER

doi:10.1038/nature11398

HIV-infected T cells are migratory vehicles for viral dissemination

Thomas T. Murooka[1], Maud Deruaz[1], Francesco Marangoni[1], Vladimir D. Vrbanac[1], Edward Seung[1], Ulrich H. von Andrian[2], Andrew M. Tager[1], Andrew D. Luster[1] & Thorsten R. Mempel[1]

After host entry through mucosal surfaces, human immunodeficiency virus-1 (HIV-1) disseminates to lymphoid tissues to establish a generalized infection of the immune system. The mechanisms by which this virus spreads among permissive target cells locally during the early stages of transmission and systemically during subsequent dissemination are not known[1]. *In vitro* studies suggest that the formation of virological synapses during stable contacts between infected and uninfected T cells greatly increases the efficiency of viral transfer[2]. It is unclear, however, whether T-cell contacts are sufficiently stable *in vivo* to allow for functional synapse formation under the conditions of perpetual cell motility in epithelial[3] and lymphoid tissues[4]. Here, using multiphoton intravital microscopy, we examine the dynamic behaviour of HIV-infected T cells in the lymph nodes of humanized mice. We find that most productively infected T cells migrate robustly, resulting in their even distribution throughout the lymph node cortex. A subset of infected cells formed multinucleated syncytia through HIV envelope-dependent cell fusion. Both uncoordinated motility of syncytia and adhesion to CD4^{+} lymph node cells led to the formation of long membrane tethers, increasing cell lengths to up to ten times that of migrating uninfected T cells. Blocking the egress of migratory T cells from the lymph nodes into efferent lymph vessels, and thus interrupting T-cell recirculation, limited HIV dissemination and strongly reduced plasma viraemia. Thus, we have found that HIV-infected T cells are motile, form syncytia and establish tethering interactions that may facilitate cell-to-cell transmission through virological synapses. Migration of T cells in lymph nodes therefore spreads infection locally, whereas their recirculation through tissues is important for efficient systemic viral spread, suggesting new molecular targets to antagonize HIV infection.

In HIV infection, lymph nodes are important sites of viral replication, in which the high local density of CD4^{+} T cells and other target cells may favour cell contact-mediated viral spread. Because infection of BLT (bone marrow/liver/thymus) humanized mice[5] replicates many hallmarks of infection in humans[6–8], we used this small animal model (Supplementary Fig. 1a) and intravital microscopy to investigate the dynamic behaviour of HIV-infected T cells in the stromal environment of lymph nodes.

We first sought to validate this mouse model for studying the effects of HIV infection on T-cell migration. The presence of both naive and memory human CD4^{+} and CD8^{+} T cells in secondary lymphoid organs of BLT mice demonstrates their general capacity to home to these sites (Fig. 1a). To determine the efficiency of this trafficking

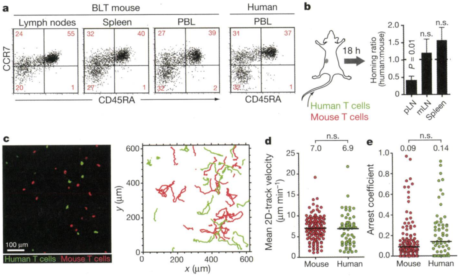


Figure 1 | Human T-cell migration in lymph nodes of BLT mice. **a**, Human CD4^{+} T cells with naive (CD45RA^{+} CCR7^{+}), central memory (CD45RA^{-}, CCR7^{+}) and effector memory (CD45RA^{-}, CCR7^{-}) phenotypes are represented in lymphoid organs and peripheral blood of BLT mice. Relative frequencies in blood are similar to humans. Plots are gated on human CD4^{+}, CD3^{+} cells. PBL, peripheral blood lymphocytes. **b**, Homing of human and mouse CD4^{+} T cells into lymphoid tissues. Data pooled from three independent experiments. Values are mean and s.e.m. mLN, mesenteric lymph nodes; pLN, peripheral lymph nodes. **c**, Multiphoton intravital micrograph of a BLT popliteal lymph node 24 h after adoptive transfer of human (green) and mouse (red) T cells. Graph on the right shows migratory tracks of each population during a 30-min recording. **d**, **e**, Mean 2D-track velocities (**d**) and arrest coefficients (**e**) of mouse and human T cells in BLT lymph nodes. Lines and numbers indicate medians. Data pooled from three recordings, two independent experiments. n.s., not significant.

---

[1]The Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. [2]Immune Disease Institute and Department of Microbiology and Immunology, Harvard Medical School, Boston, Massachusetts 02115, USA.

process we infused BLT mice with mixtures of isogenic CD4$^+$ human T cells and mouse CD4$^+$ T cells from OT-II T-cell receptor (TCR) transgenic mice. Over the course of 18 h, human T cells were half as efficient as mouse cells at entering peripheral lymph nodes, whereas their accumulation in mesenteric lymph nodes was equivalent (Fig. 1b). Thus, the interspecies molecular interactions between selectins, chemokine receptors and integrins on human T cells with their ligands expressed on mouse lymph node high endothelial venules[9] are collectively functional, albeit at a slightly reduced level.

After entering lymph nodes, T-cell migration is guided by their sensing of chemokines[10] and interactions with the paracortical fibroblastic reticulum[11], which is well preserved in BLT mice (Supplementary Fig. 1b). We compared the migration of human and mouse T cells in BLT lymph nodes by multiphoton intravital microscopy (MP-IVM). Intravenously infused human and mouse cells distributed similarly in the lymph node cortex, were indistinguishable in various aspects of motility (Fig. 1c–e, Supplementary Fig. 1c–e and Supplementary Video 1), and in this regard each resembled mouse T cells in non-humanized mouse lymph nodes[4]. We concluded that human T-cell homing to lymph nodes was efficient, and cell migration within the lymph nodes of BLT mice was, using mouse T cells as a reference, normal, allowing us to study the effects of HIV infection on this basic process.

To seed lymph nodes with HIV, we injected the footpad of BLT mice with green fluorescent protein (GFP)-expressing, CCR5-using HIV (hereafter referred to as HIV-GFP), which we derived by replacing the V3 loop region of *env* in the originally CXCR4-tropic NL4-3-IRES-GFP[12] with that of HIV BaL (Supplementary Fig. 2a, b). Subcutaneous infection reliably produced high levels of plasma viraemia (Fig. 2a) and systemic infection (Supplementary Fig. 2c). However, between two and six days after footpad infection, we detected GFP$^+$ cells in the ipsilateral popliteal lymph node only, and not in remote lymph nodes, indicating that the infection was initially contained in the primary draining lymph nodes (Fig. 2b and Supplementary Fig. 2e). The vast majority of these infected cells were resting (SSC$^{\text{low}}$), antigen-experienced (CD45RO$^+$) T cells with variable expression of CCR7. By day two, some loss of CD4 cell-surface expression was apparent[13], but this was much more pronounced at day 6, when cells had also downregulated major histocompatibility complex (MHC)-I (Supplementary Fig. 2e).

To investigate the localization of HIV-infected T cells in lymph nodes early after footpad infection, we analysed draining lymph nodes by MP-IVM two days after virus inoculation. At this time, GFP$^+$ cells were evenly distributed up to several hundred micrometres away from the subcapsular sinus (SCS), indicating that lymph-borne HIV arriving in the lymph node SCS has efficient means of infiltrating the lymph node cortex (Fig. 2c). In time-lapse recordings it became apparent that GFP-expressing lymph node cells migrated at average two-dimensional (2D) velocities of ~7 μm min^{-1} and were thus robustly motile, suggesting that their motility facilitated the local dissemination of HIV infection in lymph nodes (Fig. 2d–f and Supplementary Video 2). Because most infected T cells in early simian immunodeficiency virus (SIV)-infection of macaques are resting memory T cells[14], and the majority of infected lymph node cells in our BLT mouse model resembled antigen-experienced T cells (Fig. 2b), we compared their motility to that of *in vitro*-generated, GFP-expressing, central memory-like CD4$^+$ human T cells (hereafter termed T$_{\text{CM}}$s) that express CD62L and CCR7 and migrate to lymph nodes after adoptive transfer (Supplementary Fig. 3a–c). Compared with this reference population, the motility of HIV-infected lymph node cells was reduced (Fig. 2e, f, Supplementary Fig. 3d, e and Supplementary Videos 2 and 3), which may reflect the ability of HIV to interfere with cytoskeletal processes involved in cell migration[15,16].

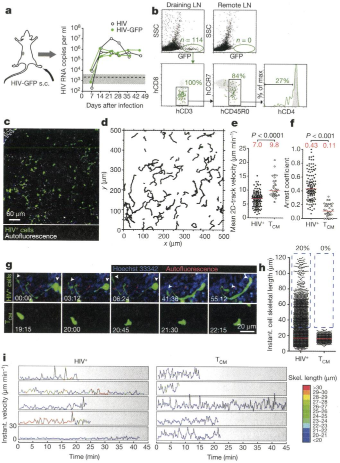


Figure 2 | *In vivo* dynamics and phenotype of HIV-infected lymph node cells. **a**, Footpad injection of BLT mice with HIV-GFP produces robust and sustained viraemia. 'HIV' is identical to HIV-GFP but lacks an IRES-GFP cassette. Similar results as shown here for five mice were obtained by other routes of infection (Supplementary Fig. 2d). Dashed line and grey-shaded area indicate mean and 95% confidence interval, respectively, of background signals obtained from plasma of uninfected mice. s.c., subcutaneous. **b**, Draining and non-draining lymph node (LN) cells two days after footpad infection with HIV-GFP. Grey dot plots and histograms show GFP$^-$ SSC$^{\text{low}}$ lymph node cells. **c**, An intravital micrograph recorded from a popliteal lymph node two days after footpad infection with HIV-GFP. **d**, Migratory tracks of GFP$^+$ lymph node cells during a 30-min recording. **e**, **f**, Mean 2D-track velocities (**e**) and arrest coefficients (**f**) of HIV$^+$ lymph node cells compared with uninfected, GFP-expressing T$_{\text{CM}}$s, recorded in lymph nodes of uninfected BLT mice. Lines and numbers indicate medians. Data on HIV-infected lymph node cells and T$_{\text{CM}}$s are representative of four and two independent experiments, respectively. **g**, MP-IVM time-lapse recordings of an HIV-infected lymph node cell (top) and an uninfected T$_{\text{CM}}$ (bottom) in BLT lymph nodes. Arrows indicate leading and trailing edge of the infected cell. Elapsed time in minutes:seconds. **h**, Instantaneous cell skeletal length of HIV-infected lymph node cells and T$_{\text{CM}}$s from recordings as shown in **g**. Lines indicate medians. Percentages indicate events >30 μm, highlighted by dashed blue box. **i**, Representative traces of infected lymph node cells and T$_{\text{CM}}$s showing instantaneous cell skeletal length (colour-coded) and instantaneous migratory velocity over time. Traces selected from 142 recorded in four movies, three independent experiments.

Although most infected lymph node cells resembled antigen-experienced T cells in terms of motility, size and shape, a subpopulation (10–20%) of cells stood out by their unusually elongated, thin and sometimes branched trailing edges. These cells reached average skeletal lengths of at times more than 100 μm, whereas T$_{\text{CM}}$s rarely exceeded 20 μm (Fig. 2g, h and Supplementary Videos 2–4). When we tracked skeletal lengths in individual cells along with their motility, we noted that some of the elongated cells continued to be motile (Fig. 2i). Similar results were obtained with SF162R3 (ref. 17), a GFP-expressing strain of HIV derived from the primary R5-tropic clone SF162 (Supplementary Fig. 4 and Supplementary Video 5).

Because it was possible that our flow cytometry analysis failed to detect these elongated cells (Fig. 2b), and that these were in fact not T cells, we wanted to determine unambiguously whether HIV-infected T cells could acquire such unusual morphological characteristics *in vivo*. Therefore, we adopted a co-culture method[18] to efficiently infect T_{CM}s with HIV-GFP *in vitro* (Fig. 3a), and transferred these cells into the footpads of BLT mice, from where they migrated to draining popliteal lymph nodes. Recipient BLT mice were pre-treated with a cocktail of antiviral drugs[6] to prevent infection of any host cells. Under these conditions, in which the T-cell identity of GFP^+ cells in popliteal lymph nodes was known, we observed the same overall reduction in motility as well as a sub-population of abnormally elongated cells at a similar frequency to after *in situ* infection (Fig. 3b–d and Supplementary Video 6). Co-injection of uninfected T_{CM}s also confirmed that these and HIV-infected T cells localized to the same regions of the lymph node and that decreased motility of HIV-infected cells did not result from their sequestration into a particular micro-environment (Supplementary Fig. 5 and Supplementary Video 6).

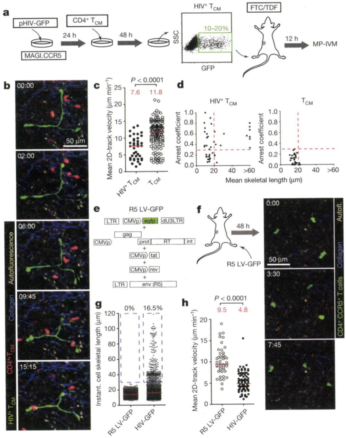


Figure 3 | HIV induces an elongated phenotype in infected T cells. **a**, Analysis of *in vitro* HIV-infected T_{CM}s in BLT lymph nodes. FTC, emtricitabine; TDF, tenofovir disoproxil fumarate. **b**, MP-IVM time-lapse series of uninfected (red) and HIV^+ (green) T_{CM}s in a BLT lymph node. Elapsed time in minutes:seconds. **c**, Mean 2D-track velocities. Red lines and numbers indicate medians. **d**, Correlation of mean cell skeletal length and arrest coefficient of individual cells. Dashed lines indicate threshold values based on measurements of uninfected T_{CM}s. Data pooled from two independent experiments. **e**, Production of an R5-tropic, GFP-expressing lentiviral vector. egfp, enhanced GFP; int, integrase; LTR, long terminal repeat; RT, reverse transcriptase. **f**, Forty-eight hours after footpad injection of 1.4×10^5 infectious units, GFP^+ cells are found in draining lymph nodes. **g**, **h**, These cells do not elongate (**g**) and do not show the reduction of cell motility of HIV-infected lymph node cells (**h**). Percentages in **g** refer to cells >30 μm, highlighted by dashed blue box. The red lines and numbers in **h** indicate medians. Data are representative of two independent experiments per three mice.

The elongated morphology could result from the infection of a T-cell population that has an intrinsic propensity to adopt such shapes. Alternatively, HIV proteins might actively induce this phenotype. To distinguish between these two possibilities, we generated a CCR5-using, HIV-derived lentiviral vector that expresses GFP but no HIV proteins in infected cells (Fig. 3e). Two days after footpad injection, a similar number of GFP^+ cells was observed in draining lymph nodes as after infection with HIV-GFP. However, these cells did not exhibit elongated shapes. They also migrated faster than HIV-GFP-infected lymph node cells and at similar speeds to T_{CM}s (Fig. 3f–h and Supplementary Video 7), further supporting the idea that HIV induces migratory deceleration of infected cells. Thus, one or several HIV proteins induces the abnormal shape and decreased motility of HIV-infected T cells.

During HIV assembly, which in migrating, polarized T cells takes place at the uropod[19], the HIV Env glycoprotein is exposed on the cell surface and can trigger the formation of intercellular adhesive contacts with CD4-expressing cells *in vitro*[20–23]. In the lymph node environment, in which T cells are highly motile, such adhesive interactions could lead to abnormal uropod elongation through tethering of infected T cells to uninfected $CD4^+$ lymph node cells. Supporting this model, we consistently observed stationary behaviour of the trailing edges of migrating T cells, often followed by their rapid retraction, perhaps reflecting sudden release of tethered uropods (Supplementary Video 8). Alternatively, cellular elongations could also result from the formation of multinucleated syncytia, which also depends on the function of Env[24]. To test whether either uropod tethering or syncytia formation are involved in the elongation of HIV-infected T cells, we generated HIV-GFP(ΔEnv), in which large parts of Env, including the CD4-binding region, are not expressed (Supplementary Fig. 6a). When *trans*-complemented with an exogenous R5 HIV envelope, the virus was as infectious as HIV-GFP, but no secondary infections occurred, as expected, because infected cells expressed no functional Env protein (Supplementary Fig. 6b–d). Notably, the elongated phenotype of infected lymph node cells was completely lost in the absence of functional Env (Fig. 4a–c and Supplementary Video 9). However, cells infected with HIV-GFP(ΔEnv) still migrated at reduced speeds (Fig. 4b, c), suggesting that HIV factors other than Env were largely responsible for the decreased T-cell motility. Cell elongation did not result from artificial interactions of Env with mouse adhesion or extracellular matrix proteins, as it was not observed in HIV-infected T_{CM}s migrating in the lymph nodes of non-humanized mice (Fig. 4d and Supplementary Fig. 7). To test whether Env interactions with CD4 specifically, or with other human ligands such as the $\alpha_4\beta_7$-integrin[25], are involved in tethering, we mutated HIV-GFP to exchange a single crucial amino acid in Env (D368R in the HBX2 consensus sequence) to abrogate CD4-binding[26] (Supplementary Fig. 8). Cells infected with HIV-GFP(D368R) showed the same loss of cell elongation as observed after Env deletion (Fig. 4d).

To examine whether some of the highly elongated T cells are part of multinucleated syncytia, we analysed infected BLT lymph nodes by histology. However, although we did find GFP-expressing cells with several nuclei (not shown), we were unable to visualize the elaborate tethers observed by MP-IVM in thick tissue sections, which may explain in part why these presumably highly labile structures have not been observed in human histological tissue samples. Therefore, we generated HIV-nGFP, in which GFP is targeted to cell nuclei through the addition of a nuclear localization signal (Supplementary Fig. 9). When we infected BLT mice with this strain, we observed by MP-IVM that the vast majority of elongated lymph node cells were multinucleated syncytia that, given the involvement of Env, probably developed through cell fusion (Fig. 4e, f). These syncytia frequently migrated with their tightly clustered nuclei moving in a coordinated fashion, but intermittently dispersed so that individual nuclei moved in different directions, yet remained connected by long membrane tethers (Supplementary Video 10). However, both fused and unfused cells also

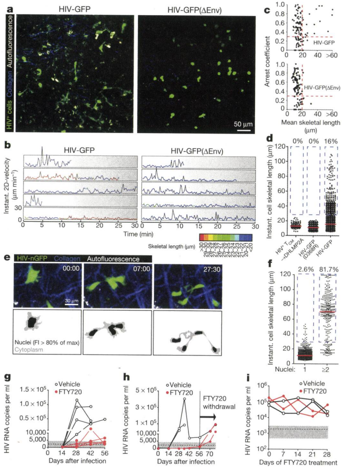


Figure 4 | HIV-infected T cells tether to other lymph node cells and form syncytia through Env, and migrate to distant tissues to disseminate infection. **a**, Intravital micrographs from lymph nodes of BLT mice injected 48 h earlier with HIV-GFP (left image) in one footpad, and with HIV-GFP(ΔEnv) (right) in the other footpad. **b**, Representative traces of infected T cells depicting instantaneous cell skeletal length (colour-coded) and migratory velocity over time. Traces selected from 281 recorded in 14 movies, 3 independent experiments. **c**, Mean cell skeletal length and arrest coefficient of individual cells infected with HIV-GFP or HIV-GFP(ΔEnv). Data are pooled from six or eight recordings, respectively, from three independent experiments. **d**, Cell lengths of HIV-GFP-infected T_{CM}s injected into natural killer cell-depleted, antibody-deficient D_HLMP2A mice, and of BLT lymph node cells infected *in situ* with HIV-GFP(D368R) or HIV-GFP. **e**, Recording of lymph node cells infected with HIV-nGFP. Bottom panels indicate border between cytoplasm and nuclei, based on an 80% maximum fluorescence intensity (FI) threshold. **f**, Cell lengths of mono- and multinucleated, HIV-infected cells. **g, h**, Viraemia in NSG BLT mice treated with FTY720 or vehicle starting at the day of s.c. HIV infection via the cheek skin. **g**, Infection with HIV-GFP. **h**, Infection with clone SF162R3. FTY720 treatment was ended at day 56. **i**, Plasma viraemia under FTY720 treatment started four weeks after intraperitoneal (i.p.) infection, when viraemia had stabilized. One representative experiment of two is shown.

formed free-ending tethers, indicating that those were caused by adhesive interactions with non-visualized CD4$^+$ cells (Supplementary Video 11). Although similar membrane tethers, such as nanotubes, can also form between T cells *in vitro* in the absence of viral infection[27], Env proteins of retroviruses such as HIV have been shown to facilitate their formation[23,28]. Adhesive tethers may promote cell-to-cell viral transfer through virological synapses, or may reflect the initiation of cell fusion, or both.

Movement of free HIV virions within and between tissues is probably restricted by the limits of diffusion, by anatomical barriers, and by soluble immune factors, such as complement and natural antibody[2]. Allowing infected T cells to retain robust motility and to serve as migratory vehicles might therefore be a viral strategy for efficient local and systemic dissemination. If long-range transport of HIV to distant tissues does rely on the trafficking of infected cells during physiological T-cell recirculation, then interfering with lymphocyte trafficking should limit systemic viral dissemination and reduce plasma viraemia.

To block T-cell egress from lymph nodes into efferent lymph vessels, we treated BLT mice with the functional sphingosine 1-phosphate receptor antagonist FTY720 (fingolimod), which resulted in profound lymphopenia (Supplementary Fig. 10a). Notably, in mice that received FTY720 at the time of infection with HIV-GFP, viral RNA in peripheral blood remained within or only slightly above the background range (Fig. 4g), whereas the drug did not show antiviral activity in cell culture (Supplementary Fig. 10b). This result was confirmed with another R5-tropic strain, SF162R3 (ref. 17; Fig. 4h). Two months after infection, viral RNA was strongly reduced, yet not absent from remote lymphoid tissues of treated animals (Supplementary Fig. 10c), suggesting that FTY720 in the longer term did not fully prevent, but only limited dissemination within the lymphoid system. Accordingly, the development of viraemia after FTY720 withdrawal following 8 weeks of treatment indicated that the release of T cells from lymphoid tissues rapidly initiated efficient systemic viral replication (Fig. 4h). Interestingly, we also observed a reduction in viral load in the draining lymph node in long-term-treated animals (Supplementary Fig. 10c). This may reflect local depletion of HIV T-cell targets, but could also indicate that FTY720 treatment reduces viral replication through a general perturbation of T-cell homeostasis or through delayed, direct antiviral activity, independently of its effect on the trafficking of HIV-infected T cells. However, FTY720 did not affect viraemia in mice with pre-established infection (Fig. 4i), similar to what was observed in simian-human immunodeficiency virus (SHIV) infection of macaques[29]. Although we cannot rule out other direct or indirect effects of FTY720 on viral replication, the observation that HIV infection is limited and high-level viraemia prevented by interfering with T-cell recirculation before, but not after, systemic viral dissemination, suggests an important role for the trafficking of infected, migratory T cells in the efficient transport of HIV from draining to remote lymphoid tissues.

In summary, we find that HIV induces a partial reduction in the motility of infected T cells, as well as their tethering to and fusion with CD4$^+$ immune cells in lymph nodes. This may reflect the need for HIV to tune T-cell migratory and interactive behaviour to meet two opposing requirements: the use of motile cells as vehicles to disseminate within and between tissues, which inherently involves short-lived cell contacts; and the formation of virological synapses and, possibly, syncytia, for cell-to-cell spread, which probably requires longer-lasting cell contacts. Reduced motility of infected T cells might favour the formation and maintenance of tethering contacts with potential HIV target cells, while still allowing for efficient trafficking to distant tissues.

METHODS SUMMARY

BLT humanized mice. Sublethally irradiated, female NOD/SCID mice were transplanted under the kidney capsule with 1 mm^3 fragments of human fetal liver and thymus, and injected intravenously with 1×10^5–5×10^5 purified CD34$^+$ human fetal liver cells. After 18–20 weeks, healthy mice with >50% human lymphoid reconstitution in peripheral blood were used for experiments within 6 weeks. In experiments that did not require popliteal lymph nodes, NOD/SCID × $gc^{-/-}$ (NSG; *gc* is also known as *Il2rg*) mice were used as graft recipients.

HIV reporter strains. The CCR5-using, GFP-expressing virus NL4-3 IRES-GFP R5 (HIV-GFP) was constructed from pBR-NL4-3-IRES-GFP-nef$^+$ (ref. 12) by replacing the V3 loop region of *env* with the corresponding sequence from the R5 clone BaL. HIV-GFP(ΔEnv) and HIV-GFP(D368R) were generated by site-directed mutagenesis. To produce infectious HIV-GFP(ΔEnv) and HIV-GFP(D368R), the BaL Env was *trans*-complemented using a helper plasmid during virus production. HIV-nGFP was generated by inserting the SV40 nuclear localization sequence amino-terminal of GFP in HIV-GFP. The lentiviral vector pHAGE-CMV-GFP was packaged with the R5 Env and standard helper plasmids to produce the CCR5-using 'R5 LV-GFP', which expresses no HIV proteins in infected cells.

MP-IVM of HIV-infected cells in lymph nodes. For *in situ*-infection of lymph node cells, BLT mice were injected into the footpad with 0.5×10^5–1.5×10^5 infectious units of recombinant HIV or R5 LV-GFP. Forty-eight hours later, the draining popliteal lymph nodes were imaged by MP-IVM, as described[4]. Alternatively, HIV-GFP-infected T_{CM}s were injected into the footpads of BLT mice pretreated with antiretroviral drugs, and imaged by MP-IVM 12 h later.

Full Methods and any associated references are available in the online version of the paper.

Received 24 February; accepted 16 July 2012.

Published online 1 August 2012.

1. Haase, A. T. Early events in sexual transmission of HIV and SIV and opportunities for interventions. *Annu. Rev. Med.* **62,** 127–139 (2011).
2. Sattentau, Q. Avoiding the void: cell-to-cell spread of human viruses. *Nature Rev. Microbiol.* **6,** 815–826 (2008).
3. Gebhardt, T. *et al.* Different patterns of peripheral migration by memory CD4$^+$ and CD8$^+$ T cells. *Nature* **477,** 216–219 (2011).
4. Mempel, T. R., Henrickson, S. E. & von Andrian, U. H. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427,** 154–159 (2004).
5. Melkus, M. W. *et al.* Humanized mice mount specific adaptive and innate immune responses to EBV and TSST-1. *Nature Med.* **12,** 1316–1322 (2006).
6. Denton, P. W. *et al.* Antiretroviral pre-exposure prophylaxis prevents vaginal transmission of HIV-1 in humanized BLT mice. *PLoS Med.* **5,** e16 (2008).
7. Sun, Z. *et al.* Intrarectal transmission, systemic infection, and CD4$^+$ T cell depletion in humanized mice infected with HIV-1. *J. Exp. Med.* **204,** 705–714 (2007).
8. Brainard, D. M. *et al.* Induction of robust cellular and humoral virus-specific adaptive immune responses in human immunodeficiency virus-infected humanized BLT mice. *J. Virol.* **83,** 7305–7321 (2009).
9. von Andrian, U. H. & Mempel, T. R. Homing and cellular traffic in lymph nodes. *Nature Rev. Immunol.* **3,** 867–878 (2003).
10. Worbs, T., Mempel, T. R., Bolter, J., von Andrian, U. H. & Forster, R. CCR7 ligands stimulate the intranodal motility of T lymphocytes *in vivo*. *J. Exp. Med.* **204,** 489–495 (2007).
11. Bajénoff, M. *et al.* Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* **25,** 989–1001 (2006).
12. Schindler, M., Munch, J. & Kirchhoff, F. Human immunodeficiency virus type 1 inhibits DNA damage-triggered apoptosis by a Nef-independent mechanism. *J. Virol.* **79,** 5489–5498 (2005).
13. Chen, B. K., Gandhi, R. T. & Baltimore, D. CD4 down-modulation during infection of human T cells with human immunodeficiency virus type 1 involves independent activities of *vpu*, *env*, and *nef*. *J. Virol.* **70,** 6044–6053 (1996).
14. Zhang, Z. *et al.* Sexual transmission and propagation of SIV and HIV in resting and activated CD4$^+$ T cells. *Science* **286,** 1353–1357 (1999).
15. Stolp, B. *et al.* HIV-1 Nef interferes with host cell motility by deregulation of Cofilin. *Cell Host Microbe* **6,** 174–186 (2009).
16. Nobile, C. *et al.* HIV-1 Nef inhibits ruffles, induces filopodia, and modulates migration of infected lymphocytes. *J. Virol.* **84,** 2282–2293 (2010).
17. Brown, A., Gartner, S., Kawano, T., Benoit, N. & Cheng-Mayer, C. HLA-A2 down-regulation on primary human macrophages infected with an M-tropic EGFP-tagged HIV-1 reporter virus. *J. Leukoc. Biol.* **78,** 675–685 (2005).
18. Casartelli, N. *et al.* Tetherin restricts productive HIV-1 cell-to-cell transmission. *PLoS Pathog.* **6,** e1000955 (2010).
19. Llewellyn, G. N., Hogue, I. B., Grover, J. R. & Ono, A. Nucleocapsid promotes localization of HIV-1 gag to uropods that participate in virological synapses between T cells. *PLoS Pathog.* **6,** e1001167 (2010).
20. Jolly, C., Kashefi, K., Hollinshead, M. & Sattentau, Q. J. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. *J. Exp. Med.* **199,** 283–293 (2004).
21. Chen, P., Hubner, W., Spinelli, M. A. & Chen, B. K. Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. *J. Virol.* **81,** 12582–12595 (2007).
22. Hübner, W. *et al.* Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. *Science* **323,** 1743–1747 (2009).
23. Rudnicka, D. *et al.* Simultaneous cell-to-cell transmission of human immunodeficiency virus to multiple targets through polysynapses. *J. Virol.* **83,** 6234–6246 (2009).
24. Sodroski, J., Goh, W. C., Rosen, C., Campbell, K. & Haseltine, W. A. Role of the HTLV-III/LAV envelope in syncytium formation and cytopathicity. *Nature* **322,** 470–474 (1986).
25. Arthos, J. *et al.* HIV-1 envelope protein binds to and signals through integrin $\alpha4\beta7$, the gut mucosal homing receptor for peripheral T cells. *Nature Immunol.* **9,** 301–309 (2008).
26. Thali, M. *et al.* Effects of changes in gp120–CD4 binding affinity on human immunodeficiency virus type 1 envelope glycoprotein function and soluble CD4 sensitivity. *J. Virol.* **65,** 5007–5012 (1991).
27. Sowinski, S. *et al.* Membrane nanotubes physically connect T cells over long distances presenting a novel route for HIV-1 transmission. *Nature Cell Biol.* **10,** 211–219 (2008).
28. Sherer, N. M. *et al.* Retroviruses can establish filopodial bridges for efficient cell-to-cell transmission. *Nature Cell Biol.* **9,** 310–315 (2007).
29. Kersh, E. N. *et al.* Evaluation of the lymphocyte trafficking drug FTY720 in SHIVSF162P3-infected rhesus macaques. *J. Antimicrob. Chemother.* **63,** 758–762 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Sodroski for the pSVIIIexE7 plasmid and A. Brown for HIV SF162R3; H. S. Shin, T. Tivey, K. Bankert and S. Tanno for technical assistance with the generation of humanized mice; A. Peixoto and D. Alvarez for management of the BL2$^+$ multiphoton microscopy facility; A. Brass, T. Allen and T. Dudek for assistance with virological techniques; and N. Elpek, M. Byrne and A. Finzi for technical assistance. Funding for this study was through National Institutes of Health (NIH) grants P01 AI0178897, R56 AI097052, R01 CA150975 and P30 AI060354, and a Platform Award from the Ragon Institute of Massachusetts General Hospital (MGH), Massachusetts Institute of Technology (MIT) and Harvard. T.T.M. was supported by the MGH ECOR Tosteson Postdoctoral Fellowship Award and NIH training grant T32 AI007387.

Author Contributions T.T.M., M.D. and T.R.M. performed all experiments. F.M. developed software for data analysis. E.S. and V.D.V. generated humanized mice. A.M.T., A.D.L. and U.H.v.A. contributed to the overall study design. T.T.M. and T.R.M. designed the experiments and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.R.M. (tmempel@mgh.harvard.edu).

LETTER

doi:10.1038/nature11419

Caspase-11 increases susceptibility to *Salmonella* infection in the absence of caspase-1

Petr Broz[1], Thomas Ruby[1], Kamila Belhocine[1], Donna M. Bouley[2], Nobuhiko Kayagaki[3], Vishva M. Dixit[3] & Denise M. Monack[1]

Inflammasomes are cytosolic multiprotein complexes assembled by intracellular nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) and they initiate innate immune responses to invading pathogens and danger signals by activating caspase-1 (ref. 1). Caspase-1 activation leads to the maturation and release of the pro-inflammatory cytokines interleukin (IL)-1β and IL-18, as well as lytic inflammatory cell death known as pyroptosis[2]. Recently, a new non-canonical inflammasome was described that activates caspase-11, a pro-inflammatory caspase required for lipopolysaccharide-induced lethality[3]. This study also highlighted that previously generated caspase-1 knockout mice lack a functional allele of *Casp11* (also known as *Casp4*), making them functionally *Casp1 Casp11* double knockouts[3–6]. Previous studies have shown that these mice are more susceptible to infections with microbial pathogens[1], including the bacterial pathogen *Salmonella enterica* serovar Typhimurium (*S. typhimurium*)[7,8], but the individual contributions of caspase-1 and caspase-11 to this phenotype are not known. Here we show that non-canonical caspase-11 activation contributes to macrophage death during *S. typhimurium* infection. Toll-like receptor 4 (TLR4)-dependent and TIR-domain-containing adaptor-inducing interferon-β (TRIF)-dependent interferon-β production is crucial for caspase-11 activation in macrophages, but is only partially required for pro-caspase-11 expression, consistent with the existence of an interferon-inducible activator of caspase-11. Furthermore, *Casp1*$^{-/-}$ mice were significantly more susceptible to infection with *S. typhimurium* than mice lacking both pro-inflammatory caspases (*Casp1*$^{-/-}$ *Casp11*$^{-/-}$). This phenotype was accompanied by higher bacterial counts, the formation of extracellular bacterial microcolonies in the infected tissue and a defect in neutrophil-mediated clearance. These results indicate that caspase-11-dependent cell death is detrimental to the host in the absence of caspase-1-mediated innate immunity, resulting in extracellular replication of a facultative intracellular bacterial pathogen.

Previous studies have shown that logarithmic phase *S. typhimurium* induce a rapid NLRC4-dependent cell death in cultured macrophages that requires the type 3 secretion system (T3SS) encoded by *Salmonella* pathogenicity island 1 (SPI-1)[9]. We previously reported that *Salmonella* grown to stationary phase (decreased SPI-1 expression) do not induce rapid NLRC4 activation, but establish themselves in an intracellular niche[10]. Intracellular *Salmonella* are detected by the inflammasome receptors NLRP3 and NLRC4, and mature IL-1β and IL-18 are released 12–17 h after infection (Supplementary Figs 1 and 2a, b)[10]. In addition, intracellular *Salmonella* induce an uncharacterized form of lytic cell death that is independent of the SPI-1 T3SS (ref. 11). To investigate host factors involved in this type of *Salmonella*-induced macrophage death, we infected macrophages deficient for specific inflammasome components with stationary phase wild-type *Salmonella*. Macrophage death, which required infection with live bacteria, did not require NLRP3 or the adaptor protein apoptosis-associated speck-like protein containing a CARD (ASC), but was partially dependent on NLRC4 (Supplementary Figs 2a and 3). Because macrophages from *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice did not release lactate dehydrogenase (LDH), we investigated whether caspase-1 and caspase-11 were activated and processed in response to intracellular *Salmonella* (Supplementary Fig. 2a). Processed caspase-1 p20 and caspase-11 p30 subunits were detected during *Salmonella* infection, indicating that caspase-11 activation correlated with cell death (Supplementary Fig. 2a). Consistently, *Casp11*$^{-/-}$ macrophages were significantly more resistant to death than wild-type macrophages (Fig. 1a), demonstrating an important role for caspase-11 in cell death caused by intracellular *Salmonella*. In contrast to *Legionella pneumophila* infections[12], intracellular growth of *Salmonella* in wild-type and *Casp11*$^{-/-}$ macrophages was not significantly different (data not shown).

Macrophage death was not totally abrogated in *Casp11*$^{-/-}$ macrophages infected with wild-type *Salmonella*, indicative of another cell death pathway (Fig. 1a and Supplementary Fig. 2c). Because NLRC4 contributed to macrophage death (Fig. 1a and Supplementary Fig. 2a), we determined whether NLRC4 activation accounted for the remaining cell death seen in *Casp11*$^{-/-}$ macrophages. The second T3SS (SPI-2), which is used by intracellular *Salmonella* to inject effector proteins into the host cell, can also inject flagellin[10], a ligand for NLRC4 (refs 13–16). Comparable levels of cell death were observed in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ and *Casp11*$^{-/-}$ macrophages infected with a ΔSPI-2 or a flagellin-deficient strain of *Salmonella* (Fig. 1a and Supplementary Fig. 2c), suggesting that wild-type *Salmonella* induce cell death through two separate pathways, one controlled by NLRC4 and the other requiring caspase-11.

Finally, we compared the levels of cell death in macrophages derived from wild-type, *Casp1*$^{-/-}$ *Casp11*$^{-/-}$, *Casp11*$^{-/-}$ or *Casp1*$^{-/-}$ (*Casp1*$^{-/-}$ *Casp11*^{tg}) mice[3] infected with *Salmonella*. Levels of cell death in *Casp1*$^{-/-}$ macrophages infected with an SPI-2-deficient strain were similar to levels seen in wild-type macrophages (Supplementary Fig. 2d), indicating that *Salmonella* strains that cannot inject flagellin exclusively engage caspase-11-dependent cell death. Consistently, *Casp11*$^{-/-}$ macrophages infected with the ΔSPI-2 strain did not die (Supplementary Fig. 2d). By contrast, wild-type *Salmonella* induced cell death by canonical (NLRC4/caspase-1) and non-canonical (caspase-11) signalling pathways (Supplementary Figs 1 and 2d).

NLRP3-mediated cytokine production triggered by non-canonical inflammasome stimuli depends on both caspase-1 and caspase-11 (ref. 3). We therefore investigated whether the NLRP3 pathway induced by intracellular *Salmonella* required both caspases. Because IL-1β and IL-18 release in response to intracellular *Salmonella* is exclusively mediated by NLRC4 and NLRP3 (Supplementary Fig. 2a, b), we studied the response to SPI-2- or flagellin-deficient *Salmonella*[10]. Cytokine maturation in response to these strains required both caspase-1 and caspase-11 (Supplementary Fig. 4). By contrast, cytokine maturation induced by wild-type *Salmonella*, which activates both NLRP3 and NLRC4, was only partially dependent on caspase-11, but absolutely required caspase-1 (Supplementary Fig. 4). These

---

[1]Department of Microbiology and Immunology, Stanford School of Medicine, Stanford University, California 94305, USA. [2]Department of Comparative Medicine Stanford School of Medicine, Stanford University, California 94305, USA. [3]Genentech Inc., South San Francisco, California 94080, USA.

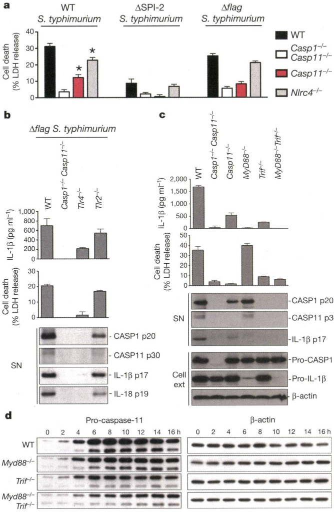


Figure 1 | Signalling through TLR4 and TRIF is required for activity of the non-canonical inflammasome pathway. **a**, LDH release from unprimed bone marrow-derived macrophages (BMDMs) infected with the indicated *S. typhimurium* strains for 17 h. **b**, **c**, IL-1β secretion, LDH release and immunoblots for processed caspase-1, caspase-11, IL-1β and IL-18 released from unprimed BMDMs infected with flagellin-deficient (Δ*flag*) *S. typhimurium* for 17 h. Ext, extract; SN, supernatant. **d**, Induction of pro-caspase-11 expression in unprimed BMDMs infected with flagellin-deficient *S. typhimurium*. Graphs show the mean and s.d. of quadruplicate wells and are representative of three (**a**–**c**) or two (**d**) independent experiments.

results indicate that intracellular *Salmonella* activate a non-canonical inflammasome similar to lipopolysaccharide/cholera toxin B and the enteric bacteria *Citrobacter rodentium*, *Vibrio cholerae* and *Escherichia coli*[3]. Formation of ASC foci (specks), a measure of NLRP3–ASC complex formation, required caspase-11 but not caspase-1 (Supplementary Fig. 5a, b), indicating that caspase-11 acts upstream of the NLRP3–ASC complex.

Stimulation of resting macrophages with lipopolysaccharide or interferon (IFN)-γ induces pro-caspase-11 expression[3,4,17]. To determine whether *Salmonella*-dependent activation of the non-canonical inflammasome was dependent on Toll-like receptor (TLR)-mediated recognition, we infected *Tlr4*$^{-/-}$ macrophages with flagellin-deficient *Salmonella* (which activate the caspase-11-dependent, non-canonical inflammasome pathway exclusively; Supplementary Figs 2 and 4). Caspase-11 activation, IL-1β secretion and cell death depended on TLR4 (Fig. 1b). Caspase-11 processing and cell death required the TLR4-dependent signalling adaptor TRIF, but not the TLR4-dependent signalling adaptor MyD88 (Fig. 1c). By contrast, IL-1β maturation was reduced in both *Myd88*$^{-/-}$ and *Trif*$^{-/-}$ (also known as *Ticam1*$^{-/-}$) macrophages, suggesting that cytokine maturation requires both adaptors. Expression of pro-IL-1β required MyD88-signalling (Supplementary Fig. 6a, b), explaining the lack of mature IL-1β release in *Myd88*$^{-/-}$ macrophages (Fig. 1c). Because cytokine production and cell death required TRIF, we measured pro-caspase-11 expression in *Myd88*$^{-/-}$, *Trif*$^{-/-}$ and *Myd88*$^{-/-}$ *Trif*$^{-/-}$ macrophages infected with *Salmonella*. Although induction of pro-caspase-11 expression was delayed in *Myd88*$^{-/-}$ and *Trif*$^{-/-}$ macrophages, the levels of pro-caspase-11 protein in *Myd88*$^{-/-}$ *Trif*$^{-/-}$ macrophages were significant (Fig. 1d and Supplementary Fig. 6a, b). Thus, pro-caspase-11 protein expression is partially dependent on TLR-signalling. However, other pathways probably contribute. Intriguingly, non-canonical inflammasomes were not activated in *Trif*$^{-/-}$ macrophages (Fig. 1c) even though significant amounts of pro-caspase-11 were present. These results indicate that caspase-11 activity requires a TRIF-dependent signal.

The adaptor protein TRIF induces nuclear factor-κB (NF-κB) activation and signals through interferon regulatory factor 3 (IRF3) to induce the expression of type-I IFNs[18]. To investigate whether type-I IFNs could be the TRIF-dependent signal required for caspase-11 activation, we compared the levels of IL-1β release, cell death and pro-caspase-11 expression in wild-type, *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ and *Irf3*$^{-/-}$ macrophages (Fig. 2a, b and Supplementary Fig. 6c). *Irf3*$^{-/-}$ macrophages were significantly impaired in their ability to initiate caspase-11-dependent IL-1β release and cell death even though significant levels of pro-caspase-11 were present, albeit at slightly reduced levels when compared with wild-type macrophages. To confirm the requirement of type-I IFN signalling for caspase-11 activation, we measured non-canonical inflammasome activation in macrophages lacking components of the type-I and type-II interferon signalling cascade. *Ifnar1*$^{-/-}$ or *STAT-1*$^{-/-}$ (also known as *Socs1*$^{-/-}$) macrophages infected with *Salmonella* did not process caspase-11 or induce non-canonical cell death (Fig. 2c), and this was not due to a lack of pro-caspase-11 expression (Fig. 2d). Macrophages lacking *Ifngr* were indistinguishable from wild-type macrophages, thereby excluding an involvement of IFN-γ. To confirm a dependency on type-I IFN signalling for caspase-11 activation, macrophages infected with flagellin-deficient *Salmonella* were treated with recombinant mouse IFN-β (Fig. 2e). Consistent with an important role for type-I IFN in caspase-11 activation, exogenous IFN-β restored cell death and caspase-11 processing in infected *Myd88*$^{-/-}$ *Trif*$^{-/-}$ but not *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ macrophages (Fig. 2e and Supplementary Fig. 7). Importantly, uninfected macrophages treated with IFN-β did not induce LDH release, confirming that IFN-β alone cannot induce non-canonical cell death in the absence of an infection. Our results reveal a previously unreported requirement for type-I IFN signalling in caspase-11 activation that is consistent with a model in which an IFN-inducible activator mediates caspase-11 activation in response to intracellular *Salmonella* (Supplementary Fig. 1).

Finally, to extend our findings to an *in vivo* setting, we infected wild-type, *Casp1*$^{-/-}$ *Casp11*$^{-/-}$, *Casp1*$^{-/-}$ (*Casp1*$^{-/-}$ *Casp11*^{tg}) and *Casp11*$^{-/-}$ mice orally with wild-type *Salmonella*. As reported previously, *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice had significantly higher tissue bacterial loads than the wild-type mice[7,8] (Fig. 3a). Surprisingly, the levels of bacteria in *Casp1*$^{-/-}$ mice were significantly higher than in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice (Fig. 3a). Interestingly, bacterial loads in *Casp11*$^{-/-}$ mice were comparable to wild-type mice in all organs examined. Unexpectedly, these results indicate that activation of pro-inflammatory caspase-11 is detrimental to the host in the absence of caspase-1. Consistent with these findings, the bacterial loads in *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$ animals, which activate caspase-11, but not caspase-1 in response to *Salmonella* (Supplementary Fig. 2), were significantly higher than in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice (Supplementary Fig. 8), essentially phenocopying *Casp1*$^{-/-}$.

Although previous studies have shown that *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice are more susceptible to many pathogens, the exact mechanism

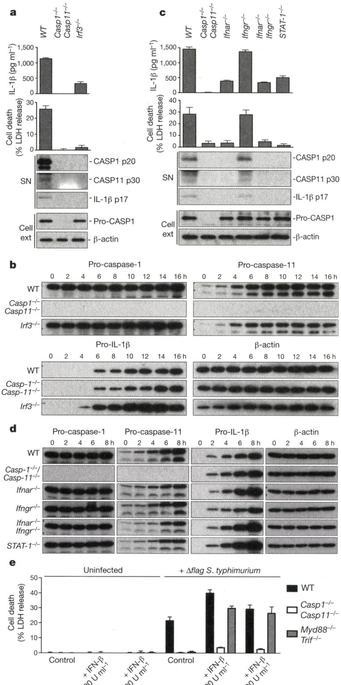


Figure 2 | Type-I IFN signalling is required for caspase-11 activation, but not for pro-caspase-11 expression. **a**, **c**, IL-1β secretion, LDH release and immunoblots for processed caspase-1 (CASP1), caspase-11 (CASP11) and IL-1β released from unprimed BMDMs infected with flagellin-deficient *S. typhimurium* for 17 h. **b**, **d**, Time course measuring pro-caspase-11 and pro-IL-1β induction in unprimed BMDMs infected with flagellin-deficient *S. typhimurium*. **e**, Cell death in BMDMs treated with recombinant mouse IFN-β or vehicle control and infected with flagellin-deficient *S. typhimurium* for 17 h. Graphs show the mean and s.d. of quadruplicate wells and are representative of three (**a**, **c**, **e**) or two (**b**, **d**) independent experiments.

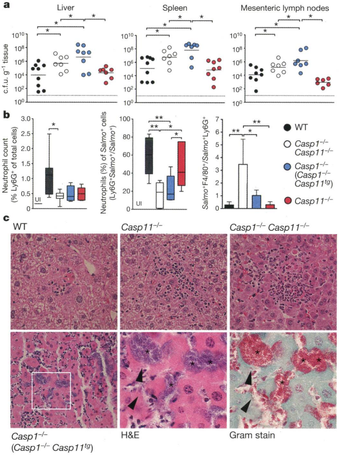


Figure 3 | $Casp1^{-/-}$ mice are more susceptible to *Salmonella* infection than $Casp1^{-/-}$ $Casp11^{-/-}$ mice. **a**, Bacterial burden in mice infected orally with green fluorescent protein $(GFP)^{+}$ wild-type *S. typhimurium* at day 4. The mean (horizontal lines) and detection limit (dashed lines) are shown. **b**, Flow-cytometry analysis of neutrophil counts $(Ly6G^{+})$, *Salmonella*-associated neutrophils $(Salmo^{+})$, and the intracellular distribution of *S. typhimurium* (GFP^{+}) in macrophages (F4/80) and neutrophils $(Ly6G^{+})$ in the spleen. UI, uninfected controls. In the box plots, lines denote the median count, and the twenty-fifth and seventy-fifth percentiles, whiskers denote maximum and minimum values. In the bar chart, the mean and s.e.m. is shown. **c**, Livers from *S. typhimurium*-infected mice showing typhoid nodules, mats of extracellular bacteria (stars) and single *Salmonella* (arrowheads). Original magnification, ×10 or ×40 (inset). H&E, haematoxylin and eosin. Graphs depict 6–10 mice per genotype and are representative of two to three independent experiments. $*P < 0.05$; $**P < 0.01$.

underlying these findings is not well understood[1,19]. Host defence against systemic *Salmonella* infection requires neutrophils, because *Salmonella* replication in the liver and spleen is exacerbated in neutropenic mice[20,21]. In addition, *Salmonella* becomes vulnerable to neutrophil-mediated clearance when it transits between host cells[22]. We therefore investigated whether caspase-1 or caspase-11-deficiency resulted in reduced neutrophil influx into the spleen. Splenic neutrophil counts ($Ly6G^{+}$ cells) were reduced in mice lacking either caspase-1 or caspase-11 in comparison with wild-type mice, with $Casp1^{-/-}$ $Casp11^{-/-}$ mice having the most significant reduction of neutrophil counts compared with wild-type animals (Fig. 3b). Because the levels of neutrophils in the mice deficient for caspase-1 or caspase-11 were very similar to one other, and did not correlate with the bacterial loads, we concluded that differences in neutrophil levels alone could not account for the higher bacterial levels in $Casp1^{-/-}$ animals (Fig. 3a). We next examined whether caspase-1 deficiency resulted in any functional defects in splenic neutrophils. Neutrophils have been previously implicated in phagocytosing and killing extracellular *Salmonella* released by pyroptotic macrophages[22]. Interestingly, the percentage of neutrophils $(Ly6G^{+})$ among all *Salmonella*-associated cells $(Salmo^{+})$ was significantly reduced in $Casp1^{-/-}$ and $Casp1^{-/-}$ $Casp11^{-/-}$ mice compared with wild-type and $Casp11^{-/-}$ mice (Fig. 3b). These data suggested that the lack of caspase-1 resulted in reduced bacterial uptake by neutrophils or reduced association with *Salmonella* during infection. However, because this reduction was observed in *Casp1*$^{-/-}$ and in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice, it did not explain the significantly higher number of colony-forming units (c.f.u.) in *Casp1*$^{-/-}$ mice (Fig. 3a).

NLRC4/caspase-1 induced lysis releases intracellular *Salmonella*, thus making them accessible to neutrophil-mediated uptake and killing[22]. Because both caspase-1 and caspase-11 can induce lysis of infected cells, we speculated that the lack of caspase-11 in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice might delay the egress of *Salmonella* from infected macrophages. Consistent with this model, we found that *Salmonella* were present to a higher degree in macrophages in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ mice than in wild-type, *Casp11*$^{-/-}$ and *Casp1*$^{-/-}$ mice (Fig. 3b). Gram stain and specific anti-*Salmonella* antibody staining of tissues showed that wild-type and *Casp11*$^{-/-}$ liver sections contained low levels of *Salmonella*, consistent with the c.f.u. data (Supplementary Fig. 9). *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ liver sections contained high levels of bacteria within cells in the sinusoids (Supplementary Fig. 9), which is consistent with our fluorescence-activated cell sorting (FACS) data indicating that a larger proportion of *Salmonella* are associated with macrophages in these mice (Fig. 3b). *Casp1*$^{-/-}$ liver sections contained mats of extracellular bacteria within typhoid nodules and expanded sinusoids (Fig. 3c and Supplementary Figs 9 and 10a). Furthermore, mice infected with *Salmonella* were injected with the membrane-impermeable antibiotic gentamicin to distinguish intracellular bacteria from extracellular bacteria. Gentamicin treatment significantly reduced bacterial counts in *Casp1*$^{-/-}$ mice, consistent with our histological finding that *Salmonella* is largely extracellular in these mice (Supplementary Fig. 10b).

We conclude that caspase-1 deficiency results in reduced neutrophil-mediated clearance of *Salmonella* released from infected macrophages, thus supporting extracellular growth of this facultative intracellular pathogen. In keeping with this observation, it has been reported that *Salmonella* rapidly replicates extracellularly in the liver of neutropenic mice[23]. This phenotype is alleviated in *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ animals, because bacterial egress from the infected macrophages is delayed. Thus, caspase-11-mediated cell death results in detrimental effects to the host in the absence of caspase-1. Our results indicate that the ability of neutrophils to phagocytose bacteria is dependent on a caspase-1-mediated function. Because *Casp1*$^{-/-}$ *Casp11*$^{-/-}$, *Il1r1*$^{-/-}$ or *Il1b*$^{-/-}$ *Il18*$^{-/-}$ mice all have similar *Salmonella* levels[8], the mechanism is probably not dependent on IL-1β and IL-18 maturation. Future studies are required to identify and characterize this function. Here we provide evidence that caspase-11-dependent cell death is exploited by *Salmonella* in the absence of caspase-1 to cause disease in the host, highlighting the need to determine whether caspase-11 activation has similar detrimental effects for the host in other infectious disease models.

METHODS SUMMARY

Mice. *Casp1*$^{-/-}$ *Casp11*$^{-/-}$ (caspase-1 knockout), *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$, *Casp11*$^{-/-}$ and *Casp1*$^{-/-}$ (*Casp1*$^{-/-}$ *Casp11*^{tg}) mice were backcrossed to C57BL/6 for at least ten generations[3,9,10]. All mouse studies were approved by the institutional animal care and use committees of Genentech and Stanford University.

Animal infections. Mice (fasted for 12 h) were inoculated orogastrically with 2.4×10^7–1×10^8 c.f.u. of wild-type or GFP$^+$ wild-type *S. typhimurium* SL1344. Tissues were collected at day 4 after infection, homogenized, and dilutions were plated on Luria–Bertani (LB) agar containing 100 μg ml^{-1} streptomycin. Bacterial numbers are expressed as c.f.u. gram^{-1} tissue.

Cell culture and infections. BMDMs were differentiated as described previously[13]. *S. typhimurium* was grown to stationary phase overnight in LB broth at 37 °C with aeration. Cells were infected at a multiplicity of infection (m.o.i.) of 100:1 and centrifuged for 15 min at 500*g* to ensure comparable adhesion of the bacteria to the cells. Gentamicin (100 μg ml^{-1}) was added 60 min after infection. Cells were washed 120 min after infection, followed by the addition of gentamicin (10 μg ml^{-1}) for the remainder of the infection. Recombinant mouse IFN-β was added 2 h after infection as indicated.

Full Methods and any associated references are available in the online version of the paper.

Received 27 March; accepted 18 July 2012.

Published online 15 August 2012.

1. Schroder, K. & Tschopp, J. The inflammasomes. *Cell* **140,** 821–832 (2010).
2. Lamkanfi, M. Emerging inflammasome effector mechanisms. *Nature Rev. Immunol.* **11,** 213–220 (2011).
3. Kayagaki, N. *et al.* Non-canonical inflammasome activation targets caspase-11. *Nature* **479,** 117–121 (2011).
4. Wang, S. *et al.* Identification and characterization of Ich-3, a member of the interleukin-1β converting enzyme (ICE)/Ced-3 family and an upstream regulator of ICE. *J. Biol. Chem.* **271,** 20580–20587 (1996).
5. Kang, S. J. *et al.* Dual role of caspase-11 in mediating activation of caspase-1 and caspase-3 under pathological conditions. *J. Cell Biol.* **149,** 613–622 (2000).
6. Kuida, K. *et al.* Altered cytokine export and apoptosis in mice deficient in interleukin-1β converting enzyme. *Science* **267,** 2000–2003 (1995).
7. Lara-Tejero, M. *et al.* Role of the caspase-1 inflammasome in *Salmonella typhimurium* pathogenesis. *J. Exp. Med.* **203,** 1407–1412 (2006).
8. Raupach, B., Peuschel, S. K., Monack, D. M. & Zychlinsky, A. Caspase-1-mediated activation of interleukin-1β (IL-1β) and IL-18 contributes to innate immune defenses against *Salmonella enterica* serovar Typhimurium infection. *Infect. Immun.* **74,** 4922–4926 (2006).
9. Mariathasan, S. *et al.* Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. *Nature* **430,** 213–218 (2004).
10. Broz, P. *et al.* Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against *Salmonella*. *J. Exp. Med.* **207,** 1745–1755 (2010).
11. Monack, D. M., Detweiler, C. S. & Falkow, S. Salmonella pathogenicity island 2-dependent macrophage death is mediated in part by the host cysteine protease caspase-1. *Cell Microbiol.* 3 825–837 (2001).
12. Akhter, A. *et al.* Caspase-11 promotes the fusion of phagosomes harboring pathogenic bacteria with lysosomes by modulating actin polymerization. *Immunity* http://dx.doi.org/10.1016/j.immuni.2012.05.001 (31 May 2012).
13. Miao, E. A. *et al.* Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1β via Ipaf. *Nature Immunol.* **7,** 569–575 (2006).
14. Franchi, L. *et al.* Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1β in *Salmonella*-infected macrophages. *Nature Immunol.* **7,** 576–582 (2006).
15. Kofoed, E. M. & Vance, R. E. Innate immune recognition of bacterial ligands by NAIPs determines inflammasome specificity. *Nature* **474,** 592–595 (2011).
16. Zhao, Y. *et al.* The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* **477,** 596–600 (2011).
17. Schauvliege, R., Vanrobaeys, J., Schotte, P. & Beyaert, R. Caspase-11 gene expression in response to lipopolysaccharide and interferon-gamma requires nuclear factor-κB and signal transducer and activator of transcription (STAT) 1. *J. Biol. Chem.* **277,** 41624–41630 (2002).
18. O"Neill, L. A. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Rev. Immunol.* **7,** 353–364 (2007).
19. Franchi, L., Munoz-Planillo, R. & Nunez, G. Sensing and reacting to microbes through the inflammasomes. *Nature Immunol.* **13,** 325–332 (2012).
20. Conlan, J. W. Critical roles of neutrophils in host defense against experimental systemic infections of mice by *Listeria monocytogenes*, *Salmonella typhimurium*, and *Yersinia enterocolitica*. *Infect. Immun.* **65,** 630–635 (1997).
21. Valdez, Y., Ferreira, R. B. & Finlay, B. B. Molecular mechanisms of Salmonella virulence and host resistance. *Curr. Top. Microbiol. Immunol.* **337,** 93–127 (2009).
22. Miao, E. A. *et al.* Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria. *Nature Immunol.* **11,** 1136–1142 (2010).
23. Conlan, J. W. Neutrophils prevent extracellular colonization of the liver microvasculature by *Salmonella typhimurium*. *Infect. Immun.* **64,** 1043–1047 (1996).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Dong, M. Wong, P. Chu and H. Matthew for technical support, G. Barton for *Tlr4*$^{-/-}$ and *Tlr2*$^{-/-}$ mice, and all members of the Monack laboratory for discussions and help with animal experiments. This work was supported by awards AI095396 and AI08972 from the National Institute of Allergy and Infectious Diseases (NIAID) to D.M.M., a Stanford Digestive Disease Center (DDC) pilot grant to P.B. and a long-term fellowship (LT000636/2009-L) from the Human Frontiers in Science Program (HFSP) to P.B.

Author Contributions P.B., K.B. and D.M.M. designed and performed the *in vitro* experiments; P.B., K.B., T.R. and D.M.M. designed and performed the *in vivo* experiments; D.M.B. performed histological analysis, N.K. and V.M.D. contributed reagents and mice; all authors analysed data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.M.M. (dmonack@stanford.edu).

CORRIGENDUM

doi:10.1038/nature11424

Corrigendum: Past extreme warming events linked to massive carbon release from thawing permafrost

Robert M. DeConto, Simone Galeotti, Mark Pagani, David Tracy, Kevin Schaefer, Tingjun Zhang, David Pollard & David J. Beerling

Nature **484**, 87–91 (2012); doi:10.1038/nature10929

In the online-only Methods of our Letter, several erroneous values appear among the parameters used to calculate peatland *D*, the average thickness of permafrost peatland deposits. Catotelm decay rate (τ) used in our equation 3 was incorrectly listed as 1,500 years instead of 15,000 years. p_c, the carbon flux from the acrotelm into the catotelm (kg m^{-2} yr^{-1}), appeared as 0.27 ± 0.19 kg m^{-2} yr^{-1} instead of 0.027 ± 0.019 kg m^{-2} yr^{-1}. Using the correct values has no effect on the calculation of peatland *D*, because the decimal point errors cancel in equation 3. In addition, peat carbon density (ρ) listed in Supplementary Table 1 was incorrect and shown as peat density instead of peat carbon density. Using the correct value for ρ in equation 3 doubles the estimated average thickness of permafrost peat deposits *D*, but does not change the calculated carbon inventories discussed in the text or in Table 2. These errors do not affect the conclusions of the Letter. All values of model parameters used in equation 3 have been corrected in the HTML and PDF versions online. The Supplementary Information has been corrected online. We thank C. J. Williams for pointing out the decimal point error associated with p_c and S. Frolking for noting the cancelling decimal point errors associated with p_c and τ, and the erroneous value of peat carbon density (ρ) and its effect on peatland *D*.

ERRATUM

doi:10.1038/nature11453

Erratum: Integrated genomic analyses of ovarian carcinoma

The Cancer Genome Atlas Research Network

Nature **474**, 609–615 (2011); doi:10.1038/nature10166

In this Letter, J. G. Vockley was incorrectly listed as J. B. Vockley. Also, several authors were inadvertently omitted from the genome sequencing centre group: C. J. Creighton, from the Human Genome Sequencing Center at the Baylor College of Medicine; R. Nichol and S. Fisher, from the Sequencing Platform at the Broad Institute; and E. Upfal, from the Department of Computer Science at Brown University. In addition, J. Samayoa, from the Institute for Computational Medicine, was inadvertently missing from the genome data analysis centre group at the Johns Hopkins University. These have been corrected in the HTML and PDF versions of the manuscript.

CORRIGENDUM

doi:10.1038/nature11515

Corrigendum: Mitochondrial DNA that escapes from autophagy causes inflammation and heart failure

Takafumi Oka, Shungo Hikoso, Osamu Yamaguchi, Manabu Taneike, Toshihiro Takeda, Takahito Tamai, Jota Oyabu, Tomokazu Murakawa, Hiroyuki Nakayama, Kazuhiko Nishida, Shizuo Akira, Akitsugu Yamamoto, Issei Komuro & Kinya Otsu

Nature **485**, 251–255 (2012); doi:10.1038/nature10992

In this Letter, several images were mistakenly switched or duplicated during preparation of the artwork. In Figs 1f and 2a, the sham-operated *Dnase2a*$^{-/-}$ and TAC-operated *Dnase2a*$^{+/+}$ mice panels were switched. In Fig. 4d, the panel showing CD3 staining for ODN2088 control-treated TAC-operated *Dnase2a*$^{+/+}$ mice (now shown correctly as black-bordered panel in Fig. 1 below) is a duplicate of that showing Ly6G staining for ODN2088-treated TAC-operated *Dnase2a*$^{-/-}$ mice. The panel showing CD45 staining for ODN2088-treated TAC-operated *Dnase2a*$^{+/+}$ (now shown correctly as blue-bordered panel in Fig. 1 below) was prepared from the original picture of ODN2088 control-treated TAC-operated *Dnase2a*$^{+/+}$. In Supplementary Fig. 4c, sham-operated *Dnase2a*$^{-/-}$ and TAC-operated *Dnase2a*$^{+/+}$ mice panels were switched. Finally, in Supplementary Fig. 10d, the panels showing CD3 and Ly6G staining for sham-operated *Tlr9*$^{+/+}$ mice were switched. These corrections do not alter any of the conclusions of this Letter, and the authors apologize for any confusion these errors may have caused.

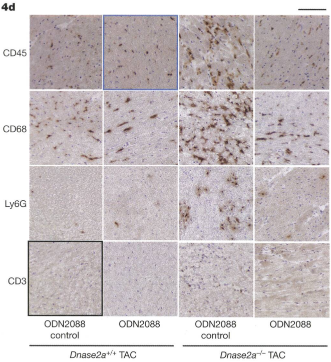


Figure 1 | This is the corrected Fig. 4d of the original Letter.

TECHNOLOGY FEATURE

CHARTING THE BRAIN'S NETWORKS

The field of connectomics is pulling neuroscience into a speedy, high-throughput lane that is generating vast amounts of data.

ALLEN INST. BRAIN SCI.

Massive stores of brain-tissue slides are providing a resource for scientists working on mapping neural networks.

BY VIVIEN MARX

Researchers seeking to understand the brain want big data. And they are getting them. Just as geneticists have moved from genes to genomes to the interacting network of factors that regulate and modify the genome, neuroscientists are going from studying single neurons to tracing how vast neuronal networks connect and interact.

"I think this is a really exciting field," says neuroscientist Moritz Helmstaedter at the Max Planck Institute for Neurobiology in Martinsried, Germany, who is working to obtain a cell-level overview of the neuronal connections — the connectome — of the mammalian cortex. "Many people are pretty ambitious about breaking the next barrier in understanding how the brain works by using this new field of connectomics."

Sceptics argue that current methods lack the power to map the massively interconnected web of around 100 billion neurons in the human brain. Even if technology can rise to the challenge, they say, it is impossible to decipher so much data.

Clay Reid, a neuroscientist at Harvard Medical School in Boston, Massachusetts, and recently appointed as a senior investigator at the Allen Institute for Brain Science in Seattle, Washington, counters detractors by pointing to recent progress in neuroscience. A few years ago, it was nearly impossible ▶

CANCER | DRUG DISCOVERY | GENETICS | IMMUNOLOGY | MICROBIOLOGY | MOLECULAR CELL BIOLOGY | NEUROSCIENCE

This month in *Nature Reviews*

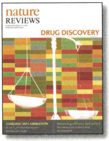
nature
REVIEWS
CANCER

Chromothripsis and cancer: causes and consequences of chromosome shattering

J. V. Forment, A. Kaidi & S. P. Jackson

Chromothripsis is an emerging phenomenon that results in chromosome rearrangements in tumour cells. This Review discusses the possible mechanisms underlying this process and its implications for cancer biology and in the clinic.

www.nature.com/reviews/cancer

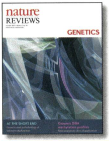


Targeting the TGFβ signalling pathway in disease

R. J. Akhurst & A. Hata

The transforming growth factor-β (TGFβ) signalling pathway is implicated in a variety of disorders, including cancer, fibrosis and inflammation, and has become an attractive target for drug development. This article presents therapeutic strategies for intervening in this pathway and discusses agents currently under investigation.

www.nature.com/reviews/drugdisc

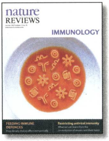


The telomere syndromes

M. Armanios & E. H. Blackburn

Telomeres have long been implicated in processes of cellular ageing. This Review discusses how a diverse range of human diseases are now known to be caused by mutations that result in defective telomere maintenance and shortened telomeres.

www.nature.com/reviews/genetics

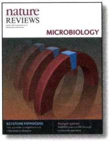


Dietary influences on intestinal immunity

M. Veldhoen & V. Brucklacher-Waldert

Dietary habits have always been associated with the robustness of the immune system. Here, the molecular links between dietary compounds and immune regulation in the intestine are discussed, with a focus on the central role of nuclear receptors.

www.nature.com/reviews/immunology

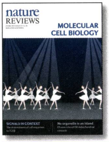


The keystone-pathogen hypothesis

G. Hajishengallis, R. P. Darveau & M. A. Curtis

In this Opinion article, Hajishengallis and colleagues propose that certain low-abundance microorganisms, termed 'keystone pathogens', can disrupt a normally benign commensal microbiota, leading to a microbial community structure that is associated with the development of disease.

www.nature.com/reviews/micro

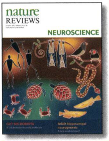


TGFβ signalling in context

J. Massagué

The basic elements of the transforming growth factor-β (TGFβ) pathway came to light more than a decade ago. Since then, the multifunctional nature and medical relevance of this pathway have relentlessly been elucidated. The old mystery of how the context determines the cellular response to TGFβ and its many malfunctions is only starting to be unravelled.

www.nature.com/reviews/molcellbio

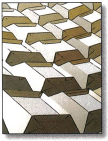


Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour

J. F. Cryan & T. G. Dinan

Gut microbiota are increasingly being recognized as influencing many aspects of human health. In this Review, Cryan and Dinan discuss rapidly emerging evidence that the gut microbiota also influence brain and behaviour and may have a role in anxiety, mood, cognition and pain.

www.nature.com/reviews/neuro



Nature Reviews Genetics **article series on Computational tools**

The scale and complexity of genetic and genomic data are ever-expanding. This series contains articles that focus on the application of software tools for the analysis, interpretation and storage of these data.

www.nature.com/nrg/series/computational

nature
REVIEWS

www.nature.com/reviews

nature publishing group npg

MPI/REF. 5

to collect data on networks of neurons. "It's not routine now but it's easier," he says. And, he adds, even without a map of the entire brain, charting just a fraction of a neural circuit is an important advance.

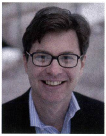
The way the neuroscience community is embracing connectomics and its big-data harvest is part of a technical and cultural shift. "We're moving away from artisan cottage-industry science and going after bigger, harder, more complex problems," says geneticist Geoffrey Duyk, who is partner and managing director at the venture investment company TPG Biotech in San Francisco, California.

Advances in high-throughput technology and automation are giving neuroscientists both speed and scale. By combining whole-brain magnetic resonance imaging and computational methods, some teams are mapping the major routes of traffic in the human brain. Other researchers are concentrating on animal models to chart the brain's neuronal circuitry on a much smaller scale, at the level of individual neurons and their projections of axons, dendrites and synaptic connections. The goal is to add the information gleaned to the animal and human connectomes, building them up as resources for understanding behaviour in health and disease.

SCALING UP

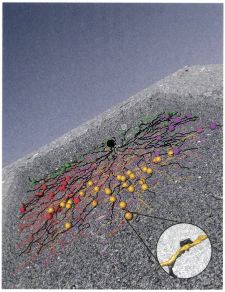
Studying the neuronal web at every possible scale demands the integration of every available method. Old-fashioned staining techniques are used alongside new methods of tissue preparation and methods taken from genomics (the study of genomes), and traditional and new approaches are used in microscopy and image analysis. Automation pipelines and computational methods are essential to handling the data — but so is skilled, manual artistry.

ALLEN INST. BRAIN SCI.



"All of us are convinced that we can get tens of thousands or maybe hundreds of thousands of connections." Clay Reid

The brain is the only organ for which the number and types of cells it contains has not been determined. Just being able to differentiate these cells from one another under a microscope is an important advance. Jeff Lichtman and his colleagues at Harvard University's Center for Brain Science in Cambridge, Massachusetts, apply light and electron microscopy to study how neural circuits change over the course of development. Using a genomic technique called Brainbow, the group was able to make specific DNA modifications in transgenic mice to label neurons fluorescently in more than 90 colours[1]. The researchers could then distinguish individual neurons in the brain's dense tangles of otherwise identical neurons. Separately, the Brainstorm Consortium, which is composed of scientists from Harvard and the Massachusetts Institute of Technology (MIT), in Cambridge, Massachusetts, and from Stanford University, California, is also working on new methods for application in connectomics, in areas such as brain-tissue preparation, imaging and image analysis.



Retinal neurons are traced to build a connectome.

Because sample preparation in neuroscience is labour and time intensive, several groups are working on improving it. Lichtman's team has developed the automatic tape-collection mechanism for ultramicrotomes (ATM), which automatically sections brain tissue and collects thousands of slices on a moving tape. The slices can then be stained and imaged on a scanning electron microscope. "It vastly improves sectioning methodology," says neuroscientist Scott Emmons, who has just set up an ATM in his lab at the Albert Einstein College of Medicine in New York. Classically, Emmons explains, sections are floated on water and then gathered manually. "They got lost, they get scrambled up, you cannot make long series," he says.

At the moment, imaging and image analysis can be only partially automated. But, for connectomics researchers to obtain the massive amounts of imaging data they want, technology will need to be expanded. Lichtman and Winfried Denk of the Max Planck Institute for Medical Research in Heidelberg, Germany, note that this approach will necessitate "the automation, even industrialization of imaging"[2].

High-throughput neuroscience reaches industrial scale at the Allen Institute for Brain Science. When a project is explored for feasibility, the institute looks for a 'brute force' method to convert it to managed, pipelined programmes with milestones and schedules, says Amy Bernard, director of structured science at the institute. This approach has been adapted from industry, says Chinh Dang, the institute's chief technology officer (see 'Neuroscience goes industrial').

Large-scale efforts at a number of other labs take on circuits in big ways. At Harvard Medical School, Reid and his colleagues have been mapping and studying neural connections in the mouse visual cortex. To do this, they first looked at a brain region *in vivo*, using two-photon microscopy — a technique that is good for imaging live tissue — and indicator molecules that show when neurons fire and release calcium. They then captured images from fixed tissue with a custom-built serial section transmission electron microscope camera array[3].

The team captured more than 3 million digital images of the mouse cortex and, in collaboration with the Pittsburgh Supercomputing Center in Pennsylvania, stitched them together into 1,200 montages comprising 10 billion pixels apiece, and aligned them in three dimensions. The computing support was key. "We would not have been able to do the work that we did last year without the help of that team," says Reid.

Denk, Helmstaedter and Kevin Briggman from the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, used a labour-intensive approach in their recent study of a mouse retinal circuit. They, too, used two-photon imaging with calcium indicators to capture brain functional patterns *in vivo* and then imaged fixed tissue to obtain ultrastructure information[4].

But for that last step they used serial block-face scanning electron microscopy, an automated technique in which a block of tissue is imaged and then the top slice is shaved off to image the next slice. With their own visualization-annotation software tools, such as KNOSSOS (www.knossostool.org), they annotated the images to show how neurons branch and interconnect.

REACHING OUT TO CROWDS

Both teams performed precise segmentation on the images they obtained: tracing the branching structure of neurons and the synapses between them. This approach has revealed links between structure and function — illuminating, for example, how the many types of neurons in the retinal circuit compute visual signals. "The way we got there was entirely segmentation by hand," says Reid, who says that the human eye remains best suited for the task, given our innate skill at recognizing patterns.

Ultimately, however, researchers such as Sebastian Seung at MIT and Dmitri Chklovskii at the Howard Hughes Institute's Janelia Farm research campus in Ashburn, Virginia, hope to teach machines how to perform segmentation.

A scientist subdivides a brain segment by hand, ready for it to be cut and mounted on slides.

ALLEN INST. BRAIN SCI.

"Image processing is currently the bottleneck," says Viren Jain, a neuroscientist at Janelia Farm.

One component of that bottleneck is the fact that circuits can be metres long. Contouring a path just one-third of a metre long would take a human annotator some 60,000 hours, or around 30 years assuming a normal work week. However, if manual reconstruction has its challenges, so too does automation, note Denk, Briggman and Helmstaedter[5,6]. Speeding up image analysis is hindered by stains that can emphasize the cell surface but obscure the visibility of synapses. And tracing a circuit is dogged by errors even with experts doing the work, which makes it hard to imagine how a computer program could do as well or better.

Ultimately, new tools are needed to determine the 'volume' of neurons — that is, when different pathways touch each other and may connect. In the meantime, Reid's Harvard team continues to collect electron microscopy data and analyse them using open-source tools such as TrakEM2 (go.nature.com/kgsalt) and the Collaborative Annotation Toolkit for Massive Amounts of Image Data (CATMAID) (www.catmaid.org). "Once the computers take over," he says, wiring diagrams will capture more connections in a circuit. "All of us are convinced that we can get tens of thousands or maybe hundreds of thousands of connections, and then it will really become circuit science."

Until then, one way of speeding up segmentation is to use crowds to sift through the data and then reconcile the many results. The software for tracing and annotating electron micrographs initially ran mainly on expensive lab computers, which limited access, says Helmstaedter. But in 2006, he and his colleagues converted KNOSSOS into distributable software and began recruiting students, of whom they now have about 200, to annotate electron micrographs at home or in the dorm. "This particular combination of being able to browse these extremely large data sets and still do it on a laptop is really so far unique," he says.

MAX PLANCK INST. NEUROBIOL.

"People are pretty ambitious about breaking the next barrier in understanding how the brain works."
Moritz Helmstaedter

The team uses different software, the redundant-skeleton consensus procedure (RESCOP), to reconcile the many annotations generated by the students, which overlap in some ways and conflict in others.

Helmstaedter and his colleagues are now looking to expand the crowd working on their data to include members of the public. To entice participation, the team has hired gaming programmers to add an element of play. "The idea is basically to fly through the brain, it has the feel of a flight simulator," he says of the prototype. Similarly, scientists at MIT have launched Wired Differently (www.wireddifferently.org) to engage crowds in neuron tracing.

In the long term, Helmstaedter hopes that automation will become so good that people are only needed for annotating complex brain regions. Having just launched his own lab, he is gearing up to pursue his goal of a cell-level connectome of the mouse cortex, which will involve tracing the connections between billions of neurons. "That is really an issue of scale," he says. "We need a factor of 100 in terms of annotation efficiency."

SMALL LABS THINK BIG

High-throughput neuroanatomy might seem to be for large labs only, but the methods they use will enable smaller ones to follow in their path, Helmstaedter says. A few smaller labs are already scaling up, using new methods to pull more data out of optical and electron microscopy images than was possible a few years ago.

In 2004, fresh from his postdoc at the University of Southern California in Los Angeles, where he studied connectivity in the rat brain, neuroscientist Hong-Wei Dong was among the first scientists to be recruited to the Allen Institute to create the Allen Brain Atlas, a map of gene expression for the entire mouse brain.

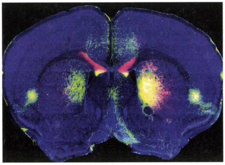
"I never thought science can be done on this kind of scale," says Dong, who has now left the Allen Institute to start his own lab at the University of California, Los Angeles (UCLA). "If I didn't have that kind of experience, I probably would have never thought of mapping the connectivity for the entire brain." At UCLA he launched iConnectome, a large-scale optical-imaging project, the aim of which is to create a three-dimensional connectome of the mouse brain. He uses optical microscopy and fluorescent markers to capture pathway information; the images are on a coarser scale than with electron microscopy but they show how brain regions interact.

Dong's methods include classic neuroanatomical techniques of surgery, preparing and mounting tissue on slides, using tracers to show neuronal inputs and outputs, and classic staining techniques such as Nissl staining, which shows the architecture of brain regions. "That is all manual labour," he says.

In neuroanatomy, tracers are usually injected into a brain region one at a time. In Dong's project, animals receive two injections in two sites, allowing scientists to examine input and output pathways concurrently, and yielding four times more data than in studies with single tracers. It also cuts down on the cost, time and number of animals needed, he says.

In the decade since his postdoc, Dong notes that scanning tissue slides has sped up dramatically, a trend he expects to continue. Microscopes can automatically scan slides one by one, each of which holds 50–100 tissue sections. Olympus, headquartered in Tokyo, Hamamatsu in Hamamatsu City, Japan, and Aperio, which was recently acquired by Leica Biosystems in Nussloch, Germany, all make high-throughput slide scanners. But speed is not everything, because analysis must be done with care. For example, Dong says that minor wrong-way tracer transport can lead to false conclusions.

And not all image-processing steps gain speed through automation. For example, tissue slices are registered automatically to a corresponding image in the Allen Reference Brain Atlas, but contrast and brightness require manual adjustment.



Tracers are used to map neural networks.

DONG LAB/UCLA

Even the process of automatic registration can be difficult, because slices differ slightly from one another, causing distortion. To match each image to the reference, Dong's team keeps enhancing the registration algorithm. "That is a difficult part, but a key part for the future if we want to evolve high-throughput analysis of the data," he says. It is also not easy to keep good computing staff, who often give in to the lure of Silicon Valley.

The three-dimensional visualization in iConnectome allows users to navigate through the brain's pathways, and layers images of tracer-labelled neurons against the backdrop of the atlas reference images, giving them a geographical context. "We put this map behind each individual brain section, and as people change the opacity, the atlas can give them a reference," Dong says. The combination of high-throughput approaches and online presentation enables brain-structure information to be shared beyond neuroanatomy circles, he says, to researchers focusing on physiology or behaviour.

While Dong finds ways to offer data-laden images to others, in his garage sit dozens of boxes of his own slides — around 6,000 of them, containing a total of 60,000 individual tissue slices from 200 rat brains. He hopes to

Neuroscience goes industrial

The Allen Institute for Brain Science in Seattle, Washington, runs efforts to create and offer online public resources about the brain, including atlases of the human, mouse and developing brains, as well as a mouse brain connectivity atlas. One of its major goals has been to combine studies of neuroanatomy and gene expression into "an integrated means of understanding the brain", says Amy Bernard, director of structured science at the institute. And that means amassing data. Now the institute is moving into broader areas, going from genes to circuits to behaviour, which means more data still.

The use of robots speeds up time-honoured neuroanatomy techniques.

ALLEN INST. BRAIN SCI.

Every week, using automated slide assembly, up to 16,000 brain sections are mounted on slides. The slides then go through *in situ* hybridization — a method used for gene-expression analysis — which is mostly automated, with six robots handling more than 190 slides per run, and two automated slide coverslippers working five slides a minute.

According to a recent tally, the institute has 1.7 million brain-tissue sections, on almost a million slides, each of which are stored both digitally and physically in filing cabinets stacked 10 drawers high throughout the microscopy suite and in off-site storage. Stacked one on top of another, the piles of slides would be 68 kilometres high.

"We keep everything, we're extremely conservative," because researchers might need to return to the data or the samples, says Bernard. All slides are kept, as are all raw data, and all images are saved and backed up. As in smaller neuroscience labs, the institute begins with a brain that is prepared for the microscope using classic neuroanatomy techniques such as histological staining. There is "nothing new or fancy there", she says, but doing it in "huge volumes" is new.

The production pipeline includes cutting tissue; mounting the samples on slides; performing *in situ* hybridization, staining, or both; scanning the images; and saving the data into an informatics pipeline, all automated steps.

Compiling their atlas of gene expression in the mouse brain generated 600 terabytes of data in four years. "This year alone we're generating 1.7 petabytes of data," says Chinh Dangh, the institute's chief technology officer. Information flow through the pipeline of processing stages is automated, as at that scale "there is no way that anyone can manually manage that kind of data", she says.

The teams also work with instrument manufacturers to adapt technology. "We're really tinkerers," says Bernard. For example, they engineered a faster slide-feed cassette for their microscopes. "You can put in 100 slides, and it will automatically load them onto the stage of the microscope and then scoot them off the stage when it is done scanning," she says.

Data infrastructure is built from the beginning. "We don't wait until the science has been done and then people start to look at the data and then build tools around the data," says Dangh. The institute uses open-source databases and tools, and this summer invited academics and independent programmers to a 'hackathon', in which participants had access to the institute's application-programming interface to develop new data-analysis tools.

image them so that he can share his work on the stria terminalis brain region with others.

Brendan Brinkman, senior product manager of the Olympus America Scientific Equipment Group, based in Center Valley, Pennsylvania, has worked with Dong and other neuroanatomists who use slide-scanning and point-scanning confocal technology. To hasten data capture in microscopy further, the company is expanding its multi-point mapping software, Brinkman says. In scanning confocal imaging, a raster scan moves across a sample. To focus only on differences in fluorescence, the scanner can jump from one area to another and capture bursts of data. "You can adjust the scanning path to make it as fast as possible," he says. Olympus has also tailored the software to let scientists capture multiple-channel fluorescent signals. "Certainly, neurobiology is the key group for this kind of technology," he says.

And connectome projects are increasing sales of slide scanners, as researchers seek the quick generation of data from large sample sets. TPG Biotech's Duyk agrees that high-throughput approaches to neuroanatomy could be a commercial boon. "It certainly creates opportunities for the life-science tools companies to push the cutting edge of their technology," he says.

For those lacking these tools, services are emerging. Earlier this year, for example, Renovo Neural, a spin-out from the Cleveland Clinic in Ohio, launched an electron microscopy service. Customers deliver samples, which the company sections with an ultramicrotome from Gatan in Pleasanton, California, then images with an automated serial block-face scanning microscope from Zeiss in Jena, Germany, to return hundreds of ultrastructure images.

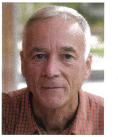
NETWORK SUCCESS

The first complete connectome was obtained for the worm *Caenorhabditis elegans* in 1986. Sydney Brenner and his colleagues at the MRC Laboratory of Molecular Biology in Cambridge, UK, completed their wiring diagram of the hermaphrodite worm nervous system by tracing images of neurons on electron micrographs by hand[7]. The hermaphrodite is one of the animal's two genders; but the male worm has proven to be tougher to pin down, neuronally speaking.

Teams tried to map the male's nervous system as Brenner's group had for the hermaphrodite, but assembling its branching structure by hand was too hard. "The field of connectomics kind of got stuck," Emmons says. Electron microscopist David Hall at the Albert Einstein College of Medicine is the custodian of material from the Brenner lab: notebooks, embedded tissue blocks, thin sections, negatives and the manually annotated *C. elegans* electron micrographs. This year, he, Emmons and colleagues revisited 5,000 of the historic images, analysing them in a new way[8].

They translated their analysis of the micrographs into a map of all the connections and their strengths in the male *C. elegans* posterior nervous system. Of the 170 neurons they studied, 144 were involved in the circuit controlling mating behaviour, allowing the team to make a link between connectivity at the cellular level and behaviour.

Previous researchers had counted synapses but had not included synapse size, Emmons says, which the team now did. Then they applied a mathematical model to use size as a proxy for the functional strength of each neural connection. The result is a map of the neural network's connectivity that includes quantitative cell-biology information. Across the connectome, they found that interaction strengths between neurons varied more than 100-fold.



"With one shot we can cover a whole worm cross-section."
Scott Emmons

ALBERT EINSTEIN COLL. MED.

Using the existing maps from Brenner's lab made neuron tracing easier. "Somebody has already gone before you and put coloured numbers on them," Emmons says. Elegance (go.nature.com/nvsfnn), the group's software tool, accelerated neuron tracing by translating mouse clicks into map coordinates as the team scrutinized the digitized micrographs.

Emmons thinks that his methods can speed up connectomics efforts in other small organisms. His group will expand its range in several ways; for example, by using their new ATM machine to automatically collect a series of slices for later imaging. Ultimately, says Emmons, using the ATM, "we're hoping for a 20,000 unbroken series, which would cover an entire worm, which has never been done before". Technology like this gives a small lab the tools to tackle large projects such as finishing the male worm's connectome, comparing it to the hermaphrodite or mapping the developing nervous system. The *C. elegans* community is "back in the connectomics business", he says.

Emmons will move on to analyse synaptic connections in the mouse brain, although, for now, his focus is on the worm. The ATM will deliver brain sections aplenty for imaging, but he is confident that his new scanning electron microscope with its bigger field of view is up to the task. "With one shot we can cover a whole worm cross-section," he says.

Applying these techniques to mammalian brains takes more than automation. The human brain has over 80 billion neurons and the mouse brain has around 70 million. And both have more densely woven webs of neurons than in the worm. "So you can't just scale up from a little lab and make a big lab and do it," Emmons says.

Big labs mapping large circuits on the single-neuron level are trying new scale-up approaches. Deciphering mammalian neural circuits is Reid's goal in his new position as senior investigator at the Allen Institute, to which he was appointed as part of the institute's ten-year US$300-million move to map connectomes and to use them to reach a broader understanding of brain function that integrates genes, circuits and behaviour.

This project, called MindScope, is an attempt to go beyond the anatomy and wiring of the brain to how things are computed in the cortex, by having scientists work side by side to study cell types, neural coding, modelling analysis and theory. "It's a dream come true," says Reid.

He will use transgenic mice to identify the different cell types in the cortex and thalamus, and will then focus on deciphering the neural coding in the visual parts of the brain using a combination of techniques: behavioural analysis, physiology, imaging with calcium indicators and electron microscopy. The results will be compiled into what he calls network anatomy, which is a wiring diagram with information piled onto it to map and understand the connectome's dizzying array of functionalities.

As the wealth of data from using different imaging modalities and from integrated large-scale projects comes in and is collected and annotated, labs large and small will still need to put their heads and computing power together for data analysis. "Astronomical amounts of connectomics data are being generated at an exponential rate; extracting meaning from it is the bottleneck that hasn't been broken," says Larry Swanson, neuroscientist at the University of Southern California and president elect of the Society for Neuroscience in Washington DC. ■

Vivien Marx *is technology editor at* Nature *and* Nature Methods.

1. Livet, J. *et al. Nature* **450,** 56–62 (2007).
2. Lichtman, J. W. & Denk, W. *Science* **334,** 618–623 (2011).
3. Bock, D. D. *et al. Nature* **471,** 177–182 (2011).
4. Briggman, K. L., Helmstaedter, M. & Denk, W. *Nature* **471,** 183–188 (2011).
5. Denk, W., Briggman, K. L. & Helmstaedter, M. *Nature Rev. Neurosci.* **13,** 351–358 (2012).
6. Helmstaedter, M., Briggman, K. L. & Denk, W. *Nature Neurosci.* **14,** 1081–1088 (2011).
7. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. Lond. B* **314,** 1–340 (1986).
8. Jarrell, T. A. *et al. Science* **337,** 437–444 (2012).

CORRECTION

The article 'Building better biobanks' (*Nature* **486,** 141–146; 2012) wrongly said that Freezerworks sells automated freezers. In fact, it makes data-management software for tracking samples held in such freezers.

nature
nature

A key thread in your scientific career

Since 1869 we've made it our mission to share the greatest developments in science and advance the discussion among the global scientific community. Whether in print, online or mobile, *Nature* is always at the forefront of science and is your forum to read, watch, listen and engage with key research, news and opinion.

How will you access the next landmark discovery?

➲ NATURE.COM/NATURE

nature publishing group npg

nature
chemistry

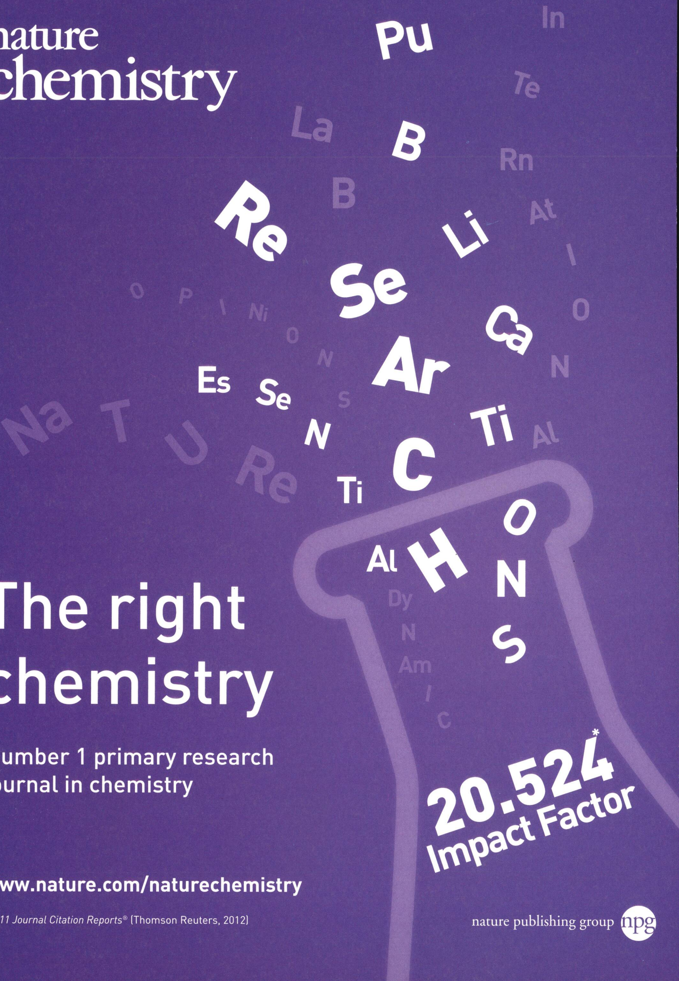
The right
chemistry

Number 1 primary research
journal in chemistry

www.nature.com/naturechemistry

*2011 *Journal Citation Reports*® (Thomson Reuters, 2012)

nature publishing group npg


Pu
In
Te
La
B
Rn
B
Re
At
Li
I
O
P
I
Ni
Se
O
Ca
O
N
N
Es
Se
N
S
Ar
Na
T
U
Re
N
Ti
Ti
Al
C
O
Al
H
N
Dy
N
S
Am
I
C
20.524*
Impact Factor

CAREERS

COLUMN Mentorship must evolve in interdisciplinary programmes **p.301**

CAREERS BLOG The latest discussions and news on research jobs **go.nature.com/z8g4a7**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

H. TRAN/IKON/CORBIS

MENTAL HEALTH

Under a cloud

Depression is rife among graduate students and postdocs. Universities are working to get them the help they need.

BY VIRGINIA GEWIN

Lauren was always a top student, but the pressures of her first year studying for a PhD in atmospheric chemistry at a UK university sent her spiralling into depression. At best, she couldn't focus on academic tasks, feeling as if her brain was "scrambled"; at worst, she couldn't get out of bed.

She developed a crippling fear of presenting her research. "Doing a PhD is such a personal thing, one that you've invested so much time in, that any criticism can feel like a direct reflection of yourself," says Lauren.

But she did something that many postgraduates do not: she got help. With counselling and medication, Lauren — a pseudonym that she uses on a blog detailing her experience (see go.nature.com/4ta9fo) — is entering the final year of her PhD. Hers is one of more than 50 stories highlighted on the website Students Against Depression, funded by the Charlie Waller Memorial Trust in Thatcham, UK. "The website aims to raise awareness that depression isn't a personal failing or weakness; it's a serious condition that requires treatment," says psychologist Denise Meyer, the website's project manager.

For early-career scientists, competing academic demands simmer in a stew of isolation, high expectations and sleeplessness that can boil over into debilitating depression, agonizing bouts of anxiety or even suicide attempts. Even if students feel that they can handle the isolation and stress of a graduate programme, extra stresses, such as problems in a relationship with an adviser or a partner, can tip them over the edge. It is important to understand the signs of depression and anxiety, know what resources are available on and off campus, and have an idea of what to expect from counselling.

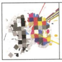
RECOGNIZING THE PROBLEM

Statistics specific to graduate students are hard to come by, but surveys[1] suggest that rates of depression have doubled among all US college students over the past 15 years, and incidence of suicidal behaviour has tripled. The best estimates are that about 10% of US college students[2], and 4% of all UK university students[3], seek treatment. And the vast majority of mental-health disorders, from schizophrenia to bipolar disorder, manifest during the teens and twenties: the college and postgraduate years.

It can be difficult for graduate students to recognize that they need help. Stress is normal — especially during training. But changes in personality or behaviour can signal problems in need of attention (see 'Signs to heed').

"Depression can have cognitive impacts that interfere with concentration and learning — things that go to the heart of being a student — making it more difficult to recognize this as depression," says Gordon Strauss, director of student mental-health services at the University of Louisville in Kentucky. "Students trying to assess whether they are depressed should ask themselves if they are able to perform research, be productive and derive satisfaction from their research."

These changes are especially tough for ▶



STRESS AND RESILIENCE
The links between adversity and mental illness. **nature.com/stress**

students who, like Lauren, have previously been high achievers. “I've seen graduate students, who have been academically successful all their lives, get to graduate school and experience academic failure for the first time — which can come as quite a shock and fuel feelings of anxiety,” says Karen Peterson, director of the office of scientific career development at the Fred Hutchinson Cancer Research Center in Seattle, Washington.

Perhaps not surprisingly, one of the main concerns for graduate students is whether they come to believe that they have chosen to specialize in the wrong field. “We see lots of students who have this concern yet can't imagine switching fields because they are so far along,” says Ronald Albucher, director of counselling and psychological services at Stanford University in California. “These are legitimate concerns that we recommend talking through with friends, family or counsellors.”

L. CICERO

“If there is a long wait for treatment, people tend not to follow through.”
Ronald Albucher

SEEKING HELP

Getting help can prove challenging. Stigma remains an issue, although more in some cultures than in others. “In Italy, very few graduate students ask for psychological help, because of the judgement that goes along with it,” says Paolo Valerio, coordinator of the Naples University Centre of Psychological Consultation, part of a campus populated by more than 90,000 students. But there are other barriers, too. “Students have competing priorities, from academic deadlines to social activities, that make it easy to put off caring for yourself — whether consciously or unconsciously,” says Daniel Eisenberg, a health-policy researcher at University of Michigan in Ann Arbor.

Students will be better prepared if they know what services are available. “When a student moves to a new place, it's a prudent idea to figure out how to get all kinds of health care — including mental-health support,” says Victor Schwartz, medical director of the Jed Foundation, a non-profit group in New York that aims to promote emotional health and prevent suicide among students. Graduate students often have access to undergraduate services; postdocs may be able to use staff resources. However, the most recent postdoc survey[4] by Sigma Xi, a scientific-research society based in Research Triangle Park, North Carolina, suggests that almost 40% of postdocs don't know whether counselling is available at their university.

“In times of crisis, it's hard to think straight,” says Julie Gold, a professional-development coach with the Office of Intramural Training and Education at the US National Institutes of Health (NIH) in Bethesda, Maryland. “I encourage incoming trainees to call the counselling centre to find out the hours and how to get an appointment.” Gold also suggests that programme directors discuss mental-health resources with students when they start and throughout the semester.

Most US and UK universities have some mental-health support on campus, but it is usually aimed at undergraduates. In mainland Europe, services are spottier: one-third of higher-education institutions don't offer counselling. And the quality of services can differ widely. “If you've seen one campus mental-health programme, you've only seen one campus mental-health programme; it's extraordinarily variable,” says Jerald Kay, chair of psychiatry at Wright State University School of Medicine in Dayton, Ohio.

COST MATTERS

In the United States, where treatment can be expensive, a student's decision to seek mental-health help “boils down to an issue of funding priorities”, says Kay. In 2006, the University of Louisville imposed a surcharge on tuition fees to expand the student health service and make psychiatric help available to more than 20,000 students. “The psychiatric services I provide are covered by the health fee so neither insurance nor parents are billed,” says Strauss.

Similarly, Roger Chalkley, an associate dean for education at Vanderbilt University in Nashville, Tennessee, has created a free, confidential service for his university's 1,100 biomedical postgraduate students.

Seeking help earlier rather than later is a good approach, say counsellors — particularly because many medications can take weeks or months to become effective. Stanford University has increased services for its 8,500 full-time graduate students and 7,000 undergraduates to ensure that people are seen more quickly. “If there is a long wait, people tend to not follow through,” says Albucher.

SYMPTOMS

Signs to heed

There are many signs of depression and anxiety — the two most common mental-health disorders among postgraduates. Anyone who experiences one or more of the following symptoms should consider seeking help:

- Inability to attend class or do research
- Difficulty concentrating
- Decreased motivation
- Increased irritability
- Sleep disturbances — insomnia, or sleep is no longer restorative
- Changes in appetite or energy levels
- Increased social withdrawal **V.G.**

One-third of US university mental-health centres place limits on the number of counselling sessions allowed. And as demand increases among UK students, university counselling services have also started offering a set number of sessions, says Patti Wallace, lead adviser for university and college counselling with the British Association for Counselling and Psychotherapy, based in Lutterworth. Once the limit is reached, counsellors will usually refer students who still need treatment to other local services.

When campus centres are overbooked or inadequate, outside services can prove useful (see ‘Where to get help’). This year, the US National Graduate Student Crisis Line (800-472-3457) has received 1,140 calls: triple the total in 2011. The rise has been fuelled in part by a mention in *The Chronicle of Higher Education* in August. “The hotline exists because of the concern that graduates students may not want to seek help locally for fear of appearing weak,” says Nick Repak, executive director of Grad Resources in Plano, Texas, which contracts and trains the Boys Town National Hotline to run the service.

COUNSELLING EXPECTATIONS

Students may be unclear on what counselling sessions will entail. Some are surprised to learn that many issues can be resolved quickly. “Many of the students I see come for only one session to talk through difficulties and receive suggestions on how better to manage stress,” says Wallace. “We aim to get people functioning effectively as quickly as possible.” Wallace focuses on offering pragmatic strategies to deal with depression, anxiety and relationship problems, but she says that each situation is unique, and it is often difficult to work out to what extent troubles are attributable to academic work or personal life.

A greater concern might be confidentiality, which is a top priority at counselling centres. “Everything students tell me is private and confidential — even if they were referred to me by an adviser or member of the faculty,” says Strauss. Counsellors will break confidence only if they feel that a patient presents imminent harm to themselves or others.

“It's important for students to understand that the counsellors offering treatment are not the agents of the university — they are there at the behest of the student's welfare,” says Kay.

However, Albucher points out that there can be a loophole: if students are still using a parent or guardian's health insurance, an explanation of services charged to the student might be sent home by the insurance company, inadvertently breaking the news about the treatment.

Postgraduates often grapple with whether to disclose mental-health issues to their advisers. “You have to be realistic that some supervisors are not necessarily the person to discuss

RESOURCES

Where to get help

In the United States:
- Grad Resources
www.gradresources.org
- The Jed Foundation
www.jedfoundation.org
- ULifeline
www.ulifeline.org
- Campus Mind Works
www.campusmindworks.org

In the United Kingdom:
- Mental Wealth
www.mentalwealthuk.com
- Student Counselling
www.student.counselling.co.uk
- Students Against Depression
www.studentsagainstdepression.org

your worries with," says Sharon Milgram, director of the NIH Office of Intramural Training and Education. Wallace suggests that postgraduates talk to a university-based counsellor about what can be gained by confiding, and what level of detail is appropriate. "If a postgraduate feels some level of disclosure would help their adviser understand their recent behaviour, it may suffice to simply say, 'I've been having problems and I'm getting help for it,'" says Wallace.

Data from the United States and the United Kingdom suggest that counselling helps the vast majority of students to get past personal problems and excel. To be successful, people must take care of themselves, notes Milgram. "Getting help," she says, "is a sign of strength, not weakness." ■

Virginia Gewin *is a freelance writer based in Portland, Oregon.*

1. Radison, K. & DiGeronimo, T. F. *College of the Overwhelmed* (Jossey-Bass, 2005).
2. Gallagher, R. P. *National Survey of Counseling Center Directors 2011* (International Assoc. of Counseling Services, 2011).
3. Royal College of Psychiatrists *Mental health of students in higher education* College Report CR166 (2011).
4. Davis, G. *Am. Scientist* **93,** (3) Supplement http://postdoc.sigmaxi.org/results/ (2005).

CORRECTION

The story 'A voice for the voiceless' (*Nature* **489,** 461–463; 2012) incorrectly stated that the US National Postdoctoral Association (NPA) is working with the Association of Public Land-grant Universities to develop a certification programme. The NPA is developing the programme on its own but will seek support and input from the university association.

COLUMN

Advising on the edge

Interdisciplinary mentorship must evolve to keep pace with innovative programmes, argues **Katherine Mackey**.

S. ELFORD

More than ever, scientists are working across disciplines. Addressing and analysing climate change, sustainable development and genomic data requires unusual and interdisciplinary approaches. As problems and solutions co-evolve in unexpected ways, skill sets must evolve too — as must mentorship. Mentors should work to bring early-career scholars up to speed on issues such as funding structures and communication styles tailored to interdisciplinary efforts — and protégés should be proactive.

Interdisciplinary work often takes root in an established department. Scholars must function in the existing framework, even when the goals, timelines and deliverables of the research differ. So when the research grows into its own programme, the first scholars to seek tenure in it will have no mentors who have navigated the same process. As early-career researchers attempt to break new ground, they must seek the few interdisciplinary mentors who can help them to find their bearings.

Young scholars should cast a wide net. Mentors from other disciplines may provide fresh perspectives, and their enthusiasm for collaborating can build an early-career researcher's confidence. But interdisciplinary scholars still need strong mentors in their own programmes, who are sensitive to extra-departmental obligations and can help them to keep pace with deadlines and degree or tenure requirements.

To ensure that their protégés get used to the culture of multiple disciplines, mentors should help students to identify and forge collaborations with other experts. This might include arranging training in different laboratories or universities, to build networks from which interdisciplinary progress can grow.

SPECIAL CHALLENGES

Logistics set interdisciplinary mentorship apart. For students, the need for course work and training in both a disciplinary department and an interdisciplinary programme can double early-stage time commitments and delay dissertation projects. It is essential for advisers and committee members to communicate with the student and each other to prioritize training objectives.

Early-career researchers on the tenure track face more serious challenges: owing to lack of precedent, institutional variability and multidisciplinary tenure committees, expectations may not be well established. How many first-authored papers should be written per year? By when should the first grant be funded? How many courses should be developed and/or taught each year? All of these milestones can vary between disciplines.

Committee or department chairs should issue annual progress reports to ensure that a scientist's research goals mesh with the department's needs. This gives the scholar a chance to adapt their research early on, to include angles or activities that the department values.

The challenge isn't just about understanding benchmarks, but also about finding ways to receive credit for unusual, difficult-to-evaluate work such as science-policy statements, decision-support systems and outreach tools. And it may not be clear whether or how extra-departmental courses, multi-authored projects or shared graduate students (who gets credit for mentoring them?) will count towards tenure requirements.

Mentors can clarify how these achievements will be 'scored' during the tenure process. And they should strive to help protégés to understand the opinions and expectations of committee members by, for example, facilitating discussions about how progress is charted.

As interdisciplinary scholars look for innovative solutions, their mentors must also develop innovative approaches. This will help the protégés to enrich their education and earn useful degrees — and chart a course towards a successful career. ■

Katherine Mackey *is a US National Science Foundation biology fellow at the Woods Hole Oceanographic Institution and the Marine Biological Laboratory in Woods Hole, Massachusetts.*

Inserm UNIVERSITE JOSEPH FOURIER SCIENCES. TECHNOLOGIE. SANTÉ CHU cea ESRF

Institut des Neurosciences GRENOBLE

The Grenoble Institute of Neuroscience (GIN) searches for its next Director

The GIN is a Research Centre considered as an Institute of excellence created in 2007 by Inserm, Joseph Fourier University - Grenoble I (UJF), CEA (Commissariat of Atomic Energy and Alternative Energies) in partnership with the Hospital of Grenoble (CHU) and the European Synchrotron Radiation Facility (ESRF).

Located in the heart of a dynamic scientific local environment (based on physics, mathematics and computer science, microelectronics, nanoscience, structural biology and chemistry), the GIN gathers teams specialized in the study of normal or/and pathological processes of the Nervous System and of the nerve-muscle ensemble. The scientific teams of the GIN are characterized by their multidisciplinary and multicultural approach of the questions they address and decipher via an emphasis onto interfaces between disciplines. The GIN provides an excellent scientific environment that allows neuroscientists and physicians to work in close proximity to each other, facilitating the path from bench to bedside. Clinical Neuroscience in Grenoble is at the origin of Deep Brain Stimulation (DBS), which found its pioneer application 25 years ago in the treatment of motor pathologies such as tremor or Parkinsonian akinesia - hypertonia. This very efficient treatment gained widespread use all over the world.

The scientific environment in Grenoble offers a facilitating ground to the development of innovative techniques. The GIN is a key link to the process of innovation with its preclinical and clinical facilities (about 6 different MRI for humans and animals up to 9.4 T, confocal and biphoton microscopy, electron microscopy, optics platform, high-resolution EEG, proteomics and transcriptomics, ..) including the presence of large European core facilities (ESRF, Grenoble antenna of EMBL - European Molecular Biology Laboratory, and PSB - Partnership in Structural Biology), the tight relations with CEA, Clinatec and Minatec innovative Centers, and with other research laboratories in Grenoble, where physics, mathematics and computer science are at the forefront.

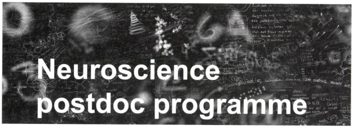
Located within the health campus of Grenoble, a few meters away from the hospital departments devoted to patients care, the GIN 6000 m2 building houses about 250 people, researchers, teachers-researchers, physicians, engineers, technicians, post-docs and PhD students along 13 teams (see description at web site http://neurosciences.ujf-grenoble.fr).

The Director will be an eminent scientist, MD, PhD or both, with a track record of excellence in fields relevant to neuroscience and documented leadership credentials in an academic research setting. He/She will serve as Chief Scientific and Executive Officer, with a mandate to expand the Centre's existing strengths, and enhance the international stature of the GIN. He/she will be expected in parallel to carry out his/her own research in the Centre, whose facilities will be made available.

Applicants should send an expression of interest, letter of motivation in support of application, research statement (not exceeding 5 pages), curriculum vitae and three letters of recommendation or names of references to claude.feuerstein@ujf-grenoble.fr at Grenoble Institute of Neuroscience – GIN, Research Centre UJF-Inserm U836-CEA-CHU, Edmond J. Safra Building, Chemin Fortuné Ferrini, Site Santé UJF at La Tronche, BP 170, 38042 Grenoble cedex 09, France.

To receive full consideration, applications should be submitted by **February 25th, 2013.**

W229464R



Neuroscience postdoc programme

Linköping University is one of Sweden's six large universities, currently enrolling 27,000 students. The university recently launched a Centre for Neurobiology, involving some 20 independent research groups from two faculties: the Faculty of Medicine and the Faculty of Science and Engineering.

The Centre for Neurobiology is now seeking Postdoctoral Fellows (2+2 years) within several neuroscience research areas: Addiction, Animal Behaviour, Electrophysiology and Circuits, Neuroimaging and Neuroengineering, Neurodegeneration, Neuroendocrinology, Neurodevelopment, Pain, and Psychiatric Functional Brain Imaging.

For more details regarding the Centre, the different research labs involved in the program, and to submit a letter-of-intent please go to: http://www.hu.liu.se/neuro?l=en. For information regarding Linköping University, please go to: http://www.hu.liu.se/?l=en, http://www.liu.se/?l=en.

Linköping University

W229341R

Massachusetts Institute of Technology

ASSISTANT PROFESSOR

MIT's Department of Brain and Cognitive Sciences anticipates making a faculty appointment at the Assistant Professor level in cognitive science or cognitive neuroscience. We seek outstanding applicants who are: (a) conducting research in areas such as language, social cognition, knowledge representation, reasoning, decision-making, or neuroeconomics; and (b) using innovative methods, such as computational modeling, functional brain imaging, behavioral genomics, and/or infant experimentation. In addition to conducting research, faculty members at MIT teach undergraduate and graduate courses, and supervise graduate and undergraduate participation in research. PhD must be completed by start date of employment and some postdoctoral training is preferred.

To apply, please submit application materials – cover letter, CV, statement of research and teaching interests and representative reprints – online at: https://academicjobsonline.org/ajo/jobs/1986. In addition, please arrange to have three letters of recommendation submitted online. Deadline for submitting applications is **November 15, 2012.**

MIT is an Affirmative Action/Equal Opportunity Employer. Women and members of underrepresented minority groups are especially encouraged to apply.

NW229036R

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Neuroscience Faculty Recruitment

The Department of Neuroscience at Columbia University is currently recruiting faculty in the neurosciences, with a primary emphasis on three research areas: 1. Cognitive and/or motor processes in awake, non-human primates; 2. Linking neural circuitry and behavior in genetically-tractable mammalian systems; and 3. Molecular, cellular and/or developmental neuroscience. In a cover letter, please direct your application to one of these three research areas. In exceptional circumstances, we will also consider applications in areas more distantly related to these three primary research themes. We encourage applications at all levels of seniority, from Assistant to full Professor.

Columbia University has an exceptionally strong and broad program in the neurosciences and aims to enhance interactions between basic and clinical research, and to link the neurosciences with a wide range of other disciplines within the University. New faculty will be affiliated with the Department of Neuroscience, with the Doctoral Program in Neurobiology and Behavior, and with a newly established Mind Brain Behavior Institute. In addition, there are numerous opportunities for interaction with other scientific departments and programs at the Medical Center and Morningside Heights campuses.

The application deadline is November 30th, 2012. Please submit applications online at https://academicjobs.columbia.edu/applicants/Central?quickFind=56746, and include a cover letter, curriculum vitae, and a statement of research interests. In addition, please arrange for three references to submit letters of recommendation.

Columbia University takes affirmative action to ensure equal opportunity.

NW228406R

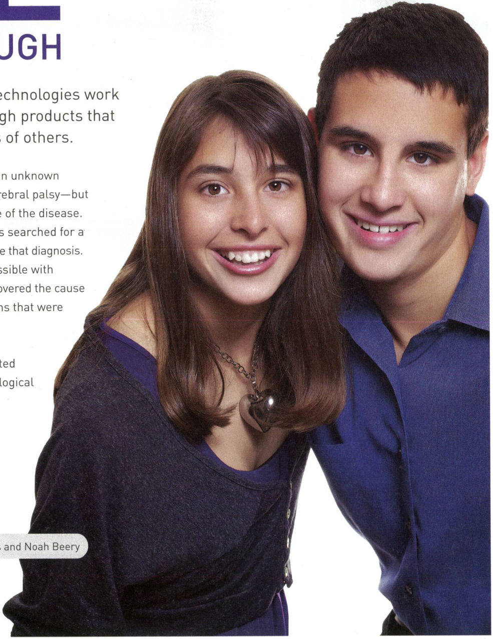
ONE
BREAKTHROUGH

Everyday, the people at Life Technologies work together to create breakthrough products that dramatically improve the lives of others.

Alexis and Noah Beery were born with an unknown condition that doctors believed to be cerebral palsy—but their symptoms were not representative of the disease. For 14 years, the twins and their parents searched for a cause, but no doctor had the tools to make that diagnosis. That is, until genetic research made possible with instruments from Life Technologies uncovered the cause of the symptoms in their DNA. Symptoms that were stopped with one simple prescription.

We are more than 10,000 people dedicated to solving some of the most difficult biological challenges of the 21st century.

Shaping Discovery.
Improving Life.



Join in to stand out. Go to **lifetechnologies.com/careers**

life technologies™

Life Technologies is an Equal Employment Opportunity/Affirmative Action employer.
©2012 Life Technologies Corporation. All rights reserved. CO03100 0712

NW229443R

ALLEN INSTITUTE
for BRAIN SCIENCE
Fueling Discovery

alleninstitute.org
brain-map.org

JOIN US AND MAKE AN IMPACT ON BRAIN RESEARCH

The Allen Institute for Brain Science is dedicated to accelerating understanding of how the human brain works in health and disease. Using a big science approach, we generate useful public resources, drive technological and analytical advances and make fundamental discoveries about the brain. Our work and efforts are intended to fuel discovery for the broader scientific community worldwide.

We are growing rapidly and seek talented scientists who have an entrepreneurial spirit and thrive as part of a collaborative, multidisciplinary team.

NEW RESEARCH PROGRAMS

Neural Coding • Cell Types • Molecular Networks

We are expanding and diversifying our programs, continuing our commitment to open science and building upon our strengths in large-scale, structured science to address fundamental questions about the brain.

Key areas of expertise include:

> Neuroscience
> Molecular Biology
> Cell Biology
> Stem Cell Biology
> Neural Coding
> Imaging
> Bioengineering
> Computational Biology
> Systems Biology
> Software Engineering

Open positions include:

> Director, Research Engineering
> Associate Director, Celluar Phenotyping
> Associate Director, Computational Biology
> Manager, Methods Development
> Scientist I, Scientist II, Senior Scientist
> Software Engineer

For more information:

Visit our website at **www.alleninstitute.org** and go to *Careers* to learn about our open positions and to apply.

We offer competitive pay, a generous benefits package, and a dynamic, intellectually challenging environment.

The Allen Institute is an equal opportunity employer.

NW22942

naturejobs

European Head Office, London naturejobs@nature.com
The Macmillan Building, 4 Crinan Street, London NI 9XW, UK
Tel: +44 (0) 20 7843 4961 Fax: +44 (0) 20 7843 4996

Global Head:
Nils Moeller,+44 (0) 20 7843 4953 n.moeller@nature.com
Southwest UK/RoW:
Emma Hare, +44 (0) 20 7843 4944 e.hare@nature.com
Northeast UK/Ireland:
Matthew Ward, +44 (0) 20 7014 4059 m.ward@nature.com
France/Switzerland/Belgium:
Julie Aramburu, +44 (0) 20 7843 4994 j.aramburu@nature.com
Scandinavia/Spain/Portugal/Italy:
Megan Ball, +44 (0) 20 7843 4973 m.ball@nature.com
North Germany/The Netherlands/Eastern Europe:
Kristin Doelling, +44 (0) 20 7843 4970 k.doelling@nature.com
South Germany/Austria:
Maria Kubalova, +44 (0) 20 7014 4084 m.kubalova@nature.com
Natureevents:
Matthew Clare, +44 (0) 20 7014 4015 m.clare@nature.com
Business Development Manager, Middle East:
Jon Giuliani, +44 (0) 20 7843 4922 j.giuliani@nature.com

USA Head Office, New York naturejobs@us.nature.com
75 Varick Street, 9th Floor, New York, NY 10013-1917
Tel: +1 800 989 7718 Fax: +1 800 989 7103

US Sales Manager:
Ken Finnegan, +1 212 726 9248 k.finnegan@us.nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751, Fax: +81 3 3267 8746

Asia-Pacific Sales & Business Development Manager:
Yuki Fujiwara, +81-3-3267-8765 y.fujiwara@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To, +852 2811 7191 g.to@natureasia.com

India
3a, 4th Floor, DLF Corporate Park, Gurgaon, 122 002, India

Advertisement Sales:
Vikas Chawla, +91 1242881057 v.chawla@nature.com

MEMORIAL SLOAN-KETTERING CANCER CENTER · ESTABLISHED 1884

Memorial Sloan-Kettering
Cancer Center

The Best Cancer Care. Anywhere.

FACULTY POSITION

Structural Biology Program

The Structural Biology Program of the Sloan-Kettering Institute (www.ski.edu) invites applications for a tenure-track faculty position at the Assistant Member level (equivalent to Assistant Professor). We are interested in individuals who have an outstanding record of research accomplishments. Areas of interest include x-ray crystallography, NMR spectroscopy, EM and optical imaging, as well as the interface of structural, chemical and computational biology. Faculty will be eligible to hold graduate school appointments in the Gerstner Sloan-Kettering Graduate School of Biomedical Sciences, the Weill Cornell Graduate School of Medical Sciences, as well as the Tri-Institutional MD/PhD Training Program.

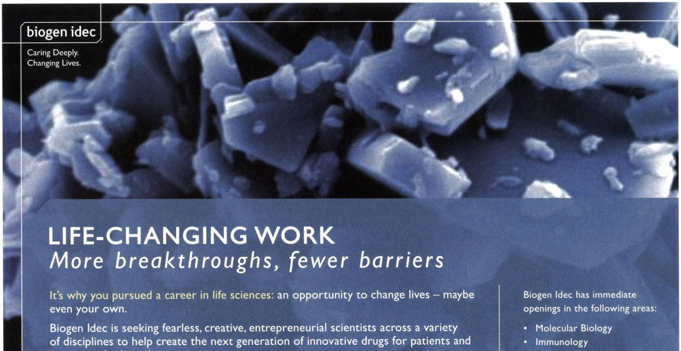
The deadline for applications is November 1, 2012. Interested candidates should visit http://facultysearch.ski.edu to access the on-line faculty application. Please visit the site as soon as possible, as it contains important information on the required application materials, including deadlines for submission of letters of reference.

Informal inquiries may be sent to **Julie Kwan** at **kwanj@mskcc.org** or to **Dr. Nikola Pavletich, Chair, Structural Biology Program** at **pavletin@mskcc.org**.

facultysearch.ski.edu

MSKCC is an equal opportunity and affirmative action employer committed to diversity and inclusion in all aspects of recruiting and employment. All qualified individuals are encouraged to apply.

NW228168R



biogen idec

Caring Deeply.
Changing Lives.

LIFE-CHANGING WORK

More breakthroughs, fewer barriers

It's why you pursued a career in life sciences: an opportunity to change lives – maybe even your own.

Biogen Idec is seeking fearless, creative, entrepreneurial scientists across a variety of disciplines to help create the next generation of innovative drugs for patients and unmet medical needs.

You'll use the latest translational and computational techniques to solve clinical problems and discover new biology. You'll do great science alongside remarkable people. It's life-changing work – and it's waiting for you at Biogen Idec.

biogenidec.com/careers

Biogen Idec has immediate openings in the following areas:

- Molecular Biology
- Immunology
- Fibrosis
- Neurobiology
- Pathology
- Cell Biology
- Antibody Engineering
- Computational Biology
- Postdoctoral Program

Biogen Idec is an Equal Opportunity/Affirmative Action Employer M/F/D/V.

NW228510R

naturejobs CAREER EXPO LONDON 2012

20 SEPTEMBER 2012

BUSINESS DESIGN CENTRE, LONDON

REVIEW FEATURE

Dear *Naturejobs* reader,

With over 1200 attendees through the door and more than 50 exhibitors, the 2012 *Naturejobs* Career Expo in London proved a hit yet again. Undoubtedly a big attraction was comic strip artist and author of PHD Comics Jorge Cham, who kicked off the conference programme with an entertaining keynote lecture. Cham shared his own experiences of postgraduate study, which he spent largely drawing cartoons. To the delight of many in the audience, he argued that procrastination is a natural part of the postgraduate process. "Listen to your inner procrastinator," he said, "relax a little." After all, Cham himself has gained international acclaim for his comics which were the direct result of his postgraduate doodlings. Cham finished off by sharing what he felt were the transferable skills anyone will gain during a PhD, whether or not you continue working in science.

Despite Cham's message that it's alright to procrastinate, there was no rest for the conference delegates who had a choice of three strands for the morning and afternoon sessions, as well as a number of workshops throughout the day. Picking up where Cham left off on the theme of transferable skills, the first session on how to survive in an economic crisis started off with a no-holds barred talk on why it's not enough to just 'be a geek'. Anna Price from Kings College London, highlighted some of the transferrable skills all scientists have, such as project management and creativity - for instance in thinking up original research questions. Stephen Tarling from the University of Southampton took a harder line, pointing out that postdocs need to broaden their knowledge base if they ever aspire to the position of vice chancellor or company CEO. Sara Williams from Cardiff University followed with tips on career development, including the importance of mentors and how to find one.

Other popular morning sessions included a talk by social media expert Nicola Osborne who gave practical advice on why scientists should embrace social networks. She pointed out that social networks encompass a lot more than the widely recognized sites like Facebook and Twitter, and include a number of networks dedicated to academics. To get the tone and content right for blogging and other social media activity, Osborne suggested seeking out some science sites which you think already do a good job, and pointed out some of her top picks (see go.nature.com/gk5asc for more). Meanwhile, Ive De Smet from the University of Nottingham shared invaluable tips on how to boost grant application success. He said it is important to seek help from more experienced people when putting a research proposal together. He also recommended looking at the strategic plans of the research councils to see what sort of areas they are particularly interested in funding.

Serial entrepreneur Ulf Grawunder, founder of NBE-Therapeutics shared his experiences of the highs and lows of stetting up your own business, and was flooded with questions by an audience keen to follow his lead. His advice was to prepare as much as possible, to create a business plan from the start, and then to go just for it: "if you are a grad student or a postdoc, and you would really like to take things into your own hands, really consider this," he urged.

In the main exhibition hall, visitors were given a chance to test out and register for the *Naturejobs* site in a new interactive iPad zone. Also new to the conference this year, a Twitter wall at the *Naturejobs* stand encouraged tweeters both at the conference and outside to get involved. Their participation formed part of the Q&A in the final session of the day, a panel discussion on the future of metrics and how best to rank institutions in the present climate, including how to take account of online activity such as social media impact, paper downloads and impact on the economy or public policy. For more information about selected talks and advice from the expo, look out for our series of articles and podcasts on the *Naturejobs* blog.

Thanks to all those who took part in the expo this year. We hope you enjoyed it and look forward to seeing you at next year's London *Naturejobs* Career Expo which will take place on September 19th 2013.

From,
The *Naturejobs* Career Expo team

Gold Sponsors:
QBRI Qatar Biomedical Research Institute
crp SANTÉ CENTRE DE RECHERCHE PUBLIC
Research in Germany Land of Ideas

Silver Sponsors:
VIB
Otsuka
BASF The Chemical Company
WORK IN DENMARK
MERCK

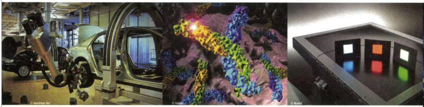
Save the date: Naturejobs Career Expo 2013 – 19 September

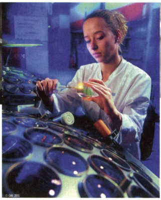
Contact Singapore
contactsingapore.sg

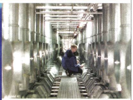
naturejobs

nature publishing group npg

Discover Research in Germany
www.research-in-germany.de







For more information on the German exhibitors at Naturejobs Career Expo 2012 visit
www.research-in-germany.de/london2012

AN INITIATIVE OF THE
Federal Ministry of Education and Research

DAAD Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

Research in Germany
Land of Ideas
www.research-in-germany.de

W228647R

QBRI
معهد قطر لبحوث الطب الحيوي
Qatar Biomedical Research Institute
Member of Qatar Foundation عضو مؤسسة قطر

EXCELLENT JOB OPPORTUNITIES IN QATAR

Qatar Biomedical Research Institute (QBRI) was established as a member of Qatar Foundation by His Highness Sheikh Hamad Bin Khalifa Al Thani, Amir of Qatar, and Her Highness Sheikha Moza bint Nasser, and opened in November 2011. QBRI is seeking suitable candidates to fill the following vacancies:

- Principal Investigators
- Senior Scientists
- Scientists
- Postdoctoral Fellows
- Research Associates

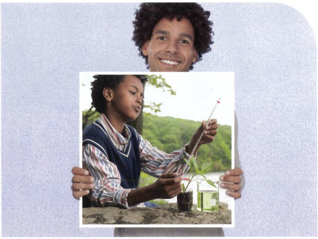
The minimum requirements for these positions are a Master's Degree or PhD in Biomedical Sciences; hands-on laboratory experience; and a strong record of publications, conference papers, reports and/or technical contributions in any of the following research fields: **Cancer, Diabetes, Neurosciences, Infectious diseases, Cardiovascular diseases, Genetic diseases, Stem cells, Bioinformatics, Metabolomics, Proteomics, Genomics, and Molecular imaging.**

Compensation will include an attractive tax-free salary and additional benefits such as furnished accommodation, annual paid leave, medical insurance, etc.
Interested applicants should send their CVs to **recruitment@qf.org.qa** with the job title as the subject of the email. Only shortlisted applicants will be notified.

About Qatar Foundation:
Qatar Foundation for Education, Science and Community Development (QF) is a private, chartered, nonprofit organization, founded in 1995 by His Highness Sheikh Hamad Bin Khalifa Al Thani, Amir of Qatar. QF is guided by the principle that a nation's greatest resource is the potential of its people, and aims to develop that potential through a network of centers devoted to progressive education, research, and community welfare, thus supporting Qatar on its journey from a carbon-based to a knowledge-based economy.

For more information please visit www.qf.org.qa

RW228590R



Research is still your favorite?

MAKE GREAT THINGS HAPPEN

Opportunities for research scientists and professionals: You have always strived to improve the patient's lives sustainably. Welcome to Merck. Contribute to exciting research projects in one of our high-growth therapeutic areas – cancer, neurodegenerative as well as autoimmune and inflammatory diseases, infertility and hormone disorders. We offer excellent development perspectives and exceptionally quick decision-making processes you will only find in mid-sized companies like ourselves. Join us and take part in shaping our diversified bio pharma business. Ready to tread new paths?

come2merck.com

MERCK

W228597R

---

challenges love solutions

We think there's a solution for every problem. That's why BASF researchers from all disciplines always work on innovations with passion. Help us, in a modern environment, to find not just products but comprehensive solutions for tomorrow's challenges. That's how we create chemistry. At BASF. Find out more now and apply at: **www.basf.com/career**

Get in touch
BASF Services Europe GmbH
Recruiting Services Europe
P.O.Box 110248
10832 Berlin, Germany
Phone: 00800-33 0000 33
E-Mail: jobs@basf.com
Career site: www.basf.com/career
Facebook: www.facebook.com/basfcareer

BASF
The Chemical Company

W228595R

Otsuka
Otsuka-people creating new products for better health worldwide

Global Staff Scientist and Clinical Research Associate Positions Available

Otsuka Pharmaceutical Co., Ltd and its group companies (http://www.otsuka.co.jp/en/) invite applications for the positions of scientist & clinical research associate to lead the global research and development of new drugs. Successful applicants will lead a team of passionate professionals who strive to advance science, technology, and medicine by developing products designed to help people live better lives.

Successful applicants will start employment as a permanent employee of Otsuka Pharmaceutical Co., Ltd or its group companies and be stationed at Otsuka facilities worldwide including Japan, USA, EU, China, and so on. The starting date is as soon as practicable in 2012.

Disease areas: **Scientists -** CNS, Circulatory, Digestive, Cancer, Infection, Immunity, Ophthalmology

Clinical Research Associates - Cancer

Experiments: **Scientists -** PhD degree with two to five years combined post-doctoral and industrial experience, preferably in the pharmaceutical industry or biotechnology sector.

Clinical Research Associates - Extensive experiences of anti-cancer drug development and expertise of regulatory registration in Europe.

Applicants should include curriculum vitae written in English, a list of publications, and a statement explaining former research/work experience. Please keep application material to no more than five pages.

Applications should be sent to: **Email: recruitRD@otsuka.jp**
(All the applications should be converted to PDF.)

JP228640R

Summit
Course
Meeting
Open Day
Virtual event
Webinar
Symposium
Congress
Lecture
Forum
Science policy event
Career fair
Seminar
Exhibition
Networking event
Space launches
International Observance
Training course
Public talk
Award Announcements
Publication date
Astronomical event
Workshop
Historical Date
Conference

DISCOVER SCIENCE EVENTS — SYNC YOUR CALENDAR TODAY.

CALENDAR VIEW NOW AVAILABLE AT NATUREEVENTS.COM

- It's easy to sync our events database with your online or mobile calendar
- Find important dates in the science calendar, not just conferences

nature publishing group npg

Communicate with Confidence

How do you speak so others want to listen?
Create an impact? Network like a pro?

VOX provides practical, imaginative training that boosts people's performance and prospects by strengthening their communication skills. We've coached researchers, academics and vice-chancellors at 25 universities, as well as staff in government departments and leading companies.

Catch us...

Catch Pete Bailie's session on Networking at the Naturejobs Expo, or email info@voxcoaching.com

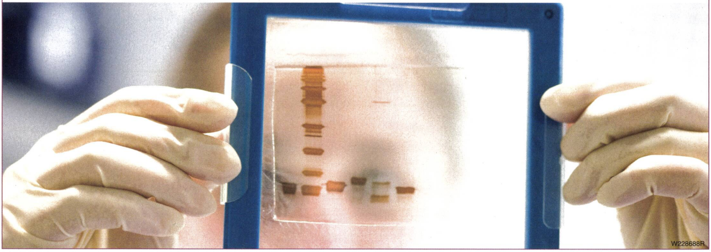
VOX Coaching
www.voxcoaching.com

U229039R

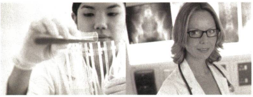
www.crp-sante.lu

crp SANTÉ
CENTRE DE RECHERCHE PUBLIC

CRP-SANTÉ: ALWAYS SEARCHING FOR MORE



W228688R



Would you like to explore career opportunities in a challenging and progressive Danish Life Science environment?

For more information please visit our website www.workindenmark.dk

UNIVERSITY OF COPENHAGEN

EURES

Technical University of Denmark DTU

WORK IN DENMARK

W229233R

Boost your career in life sciences

VIB is permanently on the look-out for enthusiastic scientists

VIB, the Flanders Institute for Biotechnology,
is excellent in distinct research areas including plant biology, neurobiology, cancer, cardiovascular medicine, inflammation, bioinformatics, systems biology, proteomics and structural biology.

www.vib.be/jobs

VIB

Currently at VIB:
- 12 postdoc and 5 PhD open positions

W229446R

naturejobs CAREER EXPO LONDON 2013

19 SEPTEMBER 2013

BUSINESS DESIGN CENTRE, LONDON

www.naturejobs.com/careerexpo

SAVE THE DATE!

naturejobs

nature publishing group npg

SIX JOINT PROFESSORS

(Monash Warwick Alliance)

The Monash Warwick Alliance represents an innovative and unique advancement in the higher education sector. Two world-class institutions are coming together to accelerate the exchange of people, ideas and information, exploiting the opportunities offered by their different geographies. This Alliance is extraordinarily well-positioned to thrive in the globalised and technology-enabled world of international higher education.

Monash University and the University of Warwick are seeking to recruit six outstanding scientists to take up joint professorial appointments.

In all cases the successful candidates will have a demonstrated world-class sustained research activity and will have shown high level leadership and a proven track record of team working and innovation. The six posts are:

- **Professor of Polymer Chemistry**
- **Professor of Polymer Materials**
- **Professor of Polymer Nanotechnology**
- **Professor of Sustainable Chemistry**
- **Professor of Sustainable Chemical Synthesis**
- **Professor of Sustainable Feedstocks and Industrial Biotechnology**

For further information and details on how apply go to: *monash.warwick.ac.uk/recruitment*

MONASH University

THE UNIVERSITY OF WARWICK

ONE WORLD. TWO UNIVERSITIES. INFINITE OPPORTUNITIES

U229440R

King Abdullah International Medical Research Center

3rd ANNUAL SCIENTIFIC FORUM

"Our Institutional Research : Moving Forward"

20th-21st November 2012

Welcome to our award-winning event where clinical decision academicians and medical researchers intermingle with the endeavor of in-depth sharing of the latest medical research findings. This, along with international speakers, various workshops, and a myriad of scientific programs and abstracts, is the recipe to our success every year. Come and be part of this innovative convention on November 20th-21st 2012.

For Abstract Submission:
E-mail: kasf_edu@ngha.med.sa

For Symposium Information:
E-mail: kaimrc_af@ngha.med.sa
Tel.: +966 1 8011111 Ext.: 49811

King Abdullah International Medical Research Center

Todays Question. Tomorrows Answer

We are an international translational health science research institution that believes in creating a knowledge-based society through our channel of extraordinary human talent.

www.kaimrc.med.sa
research1@ngha.med.sa

RW228607R

Assistant Professor

Department of Systems Biology

The Department of Systems Biology at Harvard Medical School invites applications for a tenure-track Assistant Professor position. We seek broad and careful thinkers with strong theoretical foundations (e.g. in physics, mathematics, or computer science) who have a demonstrated commitment to tackling challenging biological or medical questions. Harvard Medical School offers a unique concentration of biomedical expertise, and the Department of Systems Biology has created a supportive and interactive environment for researchers with quantitative and theoretical skills. The successful candidate will become a member of the university-wide Ph.D. Program in Systems Biology, which has an established record of attracting extraordinary graduate students.

The deadline for applications is November 30, 2012. Candidates should have a Ph.D. in a relevant discipline or an MD.

To apply, please submit a statement of research interests (~ three pages), a curriculum vitae, and the names and e-mail addresses of at least three colleagues who have agreed to write letters of recommendation to https://academicpositions.harvard.edu/postings/4336. Once you apply, your recommenders will be prompted to submit their letters on this application portal; please make sure your recommenders submit their letters before the deadline.

Candidates should have a Ph.D. in a relevant discipline or an MD.

Applications from, or nominations of, women and minority candidates are encouraged. Harvard is an affirmative action/equal opportunity employer.

NW229354R

Faculty Position in Bioengineering

The California Institute of Technology invites applications for a tenure-track faculty position in Bioengineering. Bioengineering research at Caltech focuses on the application of engineering principles to the design, analysis, construction, and manipulation of biological systems, and on the discovery and application of new engineering principles inspired by the properties of biological systems.

Applications are invited in any area of bioengineering research. Candidates with strong commitments to research and teaching excellence are encouraged to apply. We expect to make the appointment at the assistant professor level, but consideration will be given to exceptionally well-qualified applicants at the associate and full professor levels. Initial appointments at the assistant professor level are for four years, and are contingent upon completion of the Ph.D. degree.

The Bioengineering Option (http://www.be.caltech.edu) includes faculty from the Divisions of Engineering and Applied Science (http://www.eas.caltech.edu), Chemistry and Chemical Engineering (http://www.cce.caltech.edu), Biology (http://www.biology.caltech.edu), and Physics, Mathematics, and Astronomy (http://www.pma.caltech.edu).

Application Instructions

Applicants should electronically submit an application at: http://www.be.caltech.edu/search, including curriculum vitae; a statement of research interests; a statement of teaching interests; and up to three representative publications. Applicants should arrange to have four reference letters uploaded.

Application review will commence immediately and continue until the position is filled.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Division of Engineering and Applied Science
Caltech is an Equal-Opportunity/Affirmative-Action Employer.
Women, minorities, veterans, and disabled persons are encouraged to apply.

NW229367R

THE UNIVERSITY OF HONG KONG
香港大學

Founded in 1911, The University of Hong Kong is committed to the highest international standards of excellence in teaching and research, and has been at the international forefront of academic scholarship for many years. The University has a comprehensive range of study programmes and research disciplines spread across 10 faculties and about 100 sub-divisions of studies and learning. There are over 23,400 undergraduate and postgraduate students coming from 50 countries, and more than 1,800 members of academic and academic-related staff, many of whom are internationally renowned.

Tenure-Track Professor in the Department of Physics

(Ref.: 201200995)

Applications are invited for a tenure-track appointment as Professor in the Department of Physics, from as soon as possible. The post will initially be made on a three-year term basis, with the possibility of renewal subject to mutual agreement. Tenure may be offered to qualified candidates.

The position will be in any field of Physics. Existing research activities in the Department include astrophysics, condensed matter physics, high energy physics, and materials science. Candidates with excellent qualifications and commitment to high-quality research and excellence in teaching are encouraged to apply. The appointee is expected to provide academic leadership in research and teaching.

A globally competitive remuneration package commensurate with the appointee's qualifications and experience will be offered. The appointment will attract a contract-end gratuity and University contribution to a retirement benefits scheme, totalling up to 15% of basic salary, as well as housing benefits, leave, and medical benefits. At current rates, salaries tax does not exceed 15% of gross income.

For enquiries about the existing research activities and the specific job requirements, please write to Professor F.C. Zhang, Head, Department of Physics (e-mail: physhead@hku.hk). Interested applicants should submit a completed application form, together with a full C.V., a detailed publication list, a research plan, and a statement on teaching philosophy to scphy@hku.hk. Please indicate clearly "Ref.: 201200995 (Tenure-Track Professor in Department of Physics)" in the subject of the email.

Application forms (341/1111) can be obtained at http://www.hku.hk/apptunit/form-ext.doc. Further particulars can be obtained at http://jobs.hku.hk/. **Closes December 8, 2012**. The University thanks applicants for their interest, but advises that only shortlisted applicants will be notified of the application result.

The University is an equal opportunity employer and is committed to a No-Smoking Policy

JP229432R

The University of Georgia

TWO TENURE TRACK FACULTY POSITIONS IN MOLECULAR MEDICINE

The University of Georgia announces openings for two tenure track faculty positions in its new Center for Molecular Medicine (CMM). Development of the CMM complements UGA's recent thrust into human health which is marked by the opening of a new health sciences campus that includes a medical school campus operated in partnership with Georgia Health Sciences University as well as the rapidly expanding College of Public Health. The CMM, which is dedicated to better understanding the molecular and cellular basis of human disease as well as developing new disease diagnostics, therapeutics and vaccines, is currently seeking applications from talented and accomplished investigators at the Assistant or Associate Professor level. Candidates utilizing cutting-edge molecular approaches in studies of human disease etiology, tissue regeneration and repair, cancer and stem cell biology, or the application of glycoscience to medicine are especially encouraged to apply.

Applicants for the Assistant level position must have a Ph.D. degree (or equivalent), at least two years relevant postdoctoral research experience, an outstanding publication record and a well-developed research plan. Applicants for the Associate position must have a Ph.D. with a demonstrated track record of outstanding independent research productivity over a span of at least 5 years. The successful candidates will be expected to establish a vibrant, extramurally funded research program. An attractive start-up package with highly competitive conditions of employment will be offered. The University of Georgia (www.uga.edu) is a land grant/sea institution located ninety miles northeast of Atlanta (www.visitathensga.com), offers an outstanding quality of life in the Athens region (www.georgia.gov), and is an EEO/AA institution

To apply, submit a cover letter, CV, and a description of research interests to CMMSEARCH@ccrc.uga.edu electronically (indicate 'CMM Open Search' in the subject line). Applicants should arrange to have at least three letters of reference sent to: Open Search Committee Chair, c/o Ms. Karen Howard, Center for Molecular Medicine, University of Georgia, CCRC Building, 315 Riverbend Road, Athens, GA 30602-4712. Electronic letters of recommendation on letterhead are acceptable. Complete applications received by January 25th, 2013 are assured full consideration.

NW229442R

LUNDBECK FOUNDATION FELLOWSHIPS

The Lundbeck Foundation hereby invites applications for five fellowships within biomedicine and two within natural sciences which will be granted to especially promising young researchers and their research groups. Within biomedicine, the Foundation has special focus on neurology, psychiatry and allergology/immune modulation.

The fellowships are awarded for five years and each fellowship amounts to DKK 10 million (approx. Euro 1,3 million).

The subject area should be frontline basic- or applied research within the scope of the Foundation's grant strategy, which can be seen at www.lundbeckfonden.com

The fellowships may well attract Danish or foreign researchers from abroad who wish to move to Denmark and continue their research here. The call is also open for applicants at Danish universities and university hospitals.

The fellowships are intended for seven researchers who are qualified to establish or develop their own research groups within the health or natural sciences and who have received their Ph.D. degree within the last 5-7 years.

The application should include an account of the project's research plan, collaborators, budget and how the research group is envisioned to be placed within a Danish research institution. In addition, it should include a letter of intent from a resident researcher at the host institution, who makes him- or herself available as a mentor to facilitate the applicant's establishment of the research group as an integral part of the host institution. Further guidance is provided in the application form.

The application, written in English, should be sent via the Foundation's Electronic Application System for fellowships at www.lundbeckfonden.com no later than December 16, 2012.

For further information please contact Ulla Jakobsen, Science Manager at the Lundbeck Foundation, phone: +45 39 12 80 11 or at application@lundbeckfonden.com

The Lundbeck Foundation has controlling shareholdings in its subsidiaries H. Lundbeck, ALK and Falck. In addition, the Foundation manages financial investments of approx. € 1.3 billion. The Foundation supports research within the medical and natural sciences. In 2011, the Foundation had a profit after tax of approx. € 217 million and made research grants of approx. € 68 million.

Lundbeckfonden
Vestagervej 17, DK-2900 Hellerup
Tel. +45 39 12 80 00
www.lundbeckfonden.com

THE LUNDBECK FOUNDATION

W229433R

Penn

EARTH AND ENVIRONMENTAL SCIENCE

Professorship in Earth and Environmental Science

The Department of Earth and Environmental Science at the University of Pennsylvania invites applications for a tenured professorship at the Associate or Full Professor level. We seek an individual with research and teaching interests that complement and broaden our existing strengths in both Earth and Environmental Science. The successful candidate is expected to have developed an internationally recognized, externally funded, multi-disciplinary research program and will be required to actively participate in our core undergraduate and graduate teaching and in the administration of the Department. Individuals who can further increase interactions with other departments within the School of Arts and Sciences are strongly encouraged to apply. Further information about the Department may be sought at www.sas.upenn.edu/earth/.

Applicants apply online at facultysearches.provost.upenn.edu/applicants/Central?quickFind=51104 with a cover letter, CV, statements of research and teaching interests, and 3 publications. The Search Committee will begin to evaluate applications on November 15, 2012 and will continue until the position is filled. The University of Pennsylvania is an affirmative action/equal opportunity employer and is strongly committed to establishing a diverse faculty: http://www.upenn.edu/almanac/volumes/v58/n02/diversityplan.html.

NW229064R

Academic Careers Opportunties

THE UNIVERSITY OF CHICAGO

Position Title: Instructor-Professor

Req # 01475

The Department of Biochemistry and Molecular Biology at The University of Chicago invites applications for positions on its faculty at any rank. We are interested in scientists with outstanding accomplishments in and potential for fundamental research in biochemistry and/or molecular biophysics, broadly defined, and a zest for teaching. The interests of these individuals should synergize with the research in this department and other relevant biomedical programs. Generous start-up support will be provided to prime a robust research program.

Applications must include a curriculum vitae that details all publications and accomplishments, as well as plans for future research, and a teaching statement. Applications must be submitted on-line by December 14, 2012 to academiccareers.uchicago.edu/applicants/Central?quickFind=52517. Those applicants for appointment as Assistant Professor should arrange for three letters of reference to be sent to the Search Committee, Department of Biochemistry and Molecular Biology, 929 E. 57th Street, Room W225, Chicago IL 60637 or to BMB@bsd.uchicago.edu.

The University of Chicago is an Affirmative Action / Equal Opportunity Employer.

NW229449R

Wellcome Trust Clinical PhD Programmes

Establishing *innovative* PhD programmes for clinicians

Lung cancer cell. *Anne Weston/Wellcome Images*

Applications to establish PhD programmes for clinicians are invited from academic institutions in the UK and Republic of Ireland.

Programmes will provide high-quality research training opportunities that will appeal to clinicians from across the range of specialties.

Preliminary applications should be submitted by **10 December 2012**.

Information on how to apply is available at **www.wellcome.ac.uk/phdclinicians**

wellcometrust

U229395

Wellcome Trust Four-year PhD Programmes

Providing *high-quality* biomedical research training

Applications to establish new Four-year PhD Programmes are invited from academic institutions in the UK and Republic of Ireland.

Preliminary applications should be submitted by **10 December 2012**.

Information on how to apply is available at **www.wellcome.ac.uk/phdprogrammes**

Zebrafish embryo. *Annie Cavanagh/Wellcome Images*

wellcometrust

U229477R

THE SAINSBURY LABORATORY

Head of Proteomics

at The Sainsbury Laboratory, Norwich, UK

The Sainsbury Laboratory (TSL) is seeking a highly motivated individual to lead proteomics activities at the Laboratory. This is an exciting opportunity for the successful talented candidate to contribute to and enhance the world-leading scientific activities of the Laboratory.

The post holder will be responsible for managing a team providing proteomics support to TSL scientists. The key functions of the role will be to deliver proteomics and training to TSL scientists, to work closely with TSL Group Leaders to apply modern proteomic techniques, and to develop cutting-edge methodologies as opportunities arise. The post holder will have direct access to world-class facilities available at TSL and across collaborative institutes within the Norwich Research Park. A joint proteomics platform was established in 2000, which includes: Thermo Scientific LTQ-Orbitrap; Waters Synapt G2; Waters TQ-S; Bruker UltraFlex MALDI-ToF/ToF; Bruker MicroTOF, and Applied Biosystems QTrap 4000.

The applicant must have extensive hands-on experience of proteomics and must have a PhD degree in any field of research. Previous experience in the analysis of post-translational modifications of proteins is an advantage.

Salary will be within the UEA Research and Analogous Staff Grade 8 scale at between £38,140 and £48,246 pa; the appointment level will reflect qualifications, skills, knowledge and achievements. The position is available for two years in the first instance with possibility of renewal based on performance. Applicants should provide a CV, including the names and contact details of two or more referees, and a covering letter addressing the selection criteria. Please send formal applications, quoting the reference number below, either by e-mail to HR@tsl.ac.uk or by post to Kim Blanchflower, HR Manager, The Sainsbury Laboratory, Norwich Research Park, Colney, Norwich, NR4 7UH, UK.

Applications will be reviewed from 4th November 2012 and the position will remain open until filled.

The Sainsbury Laboratory (TSL) is a world leader in plant and microbial science and is dedicated to making fundamental discoveries in the science of plant-microbe interactions. TSL has expanded its scientific mission not only to continue providing fundamental biological insights into plant-pathogen interactions, but also to deliver novel, genomics-based solutions that will significantly reduce losses from major crop diseases, especially in developing countries.

For more information please visit http://www.tsl.ac.uk

Ref: HoProt/2012

U229352R

Kumamoto University

Priority Organization for Innovation and Excellence

Associate Professor

in HIV/AIDS Research

Priority Organization for Innovation and Excellence at Kumamoto University invites applications for a tenure-track Associate Professor position in HIV/AIDS research. The position is available beginning January 1, 2013, or as soon as possible thereafter (no later than February 1). Appointment at this rank is for periods up to five years.

We seek a competitive young scientist who has an outstanding record of research accomplishments and potential in HIV/AIDS research. Candidates must be 42 years old or younger and within ten years of receiving the Ph.D. degree at the time of employment. Interested applicants are encouraged to submit the following application materials: (1) a reference letter; (2) curriculum vitae; (3) a list of accomplishments; (4) a bibliography of the candidate's most significant publications and two original/photocopies of reprints for 15 publications (mainly from recent five years); (5) a statement of research and teaching accomplishments; (6) a statement of research interests and teaching philosophy; and (7) Word files of application documents listed in (2)-(6) on a recordable media (*e.g.* CD-R) to:

Mr. Koji Hamada
General Affairs Section-Research Center
Administrative Office for Life Sciences
Kumamoto University
2-2-1 Honjo, Chuo-ku, Kumamoto 860-0811, Japan
TEL: +81-96-373-6637
E-mail: iys-senter@jimu.kumamoto-u.ac.jp

Closing Date: **Friday, November 16, 2012**

Full information and application forms for (2)-(6) are available at:
http://www.caids.kumamoto-u.ac.jp/aidsnew/englishpage/
Center for AIDS Research, Kumamoto University, Japan

Kumamoto University is committed to Gender Equality and seeks to achieve equality of opportunity based on *the Basic Law for a Gender-Equal Society.*

JP229324R

Imperial College London

MRC | MRC Centre for Molecular Bacteriology and Infection

Division of Cell & Molecular Biology
Department of Life Sciences
MRC Centre for Molecular Bacteriology and Infection

Lectureships in Molecular Microbiology

Salary: £44,150 - £49,200 p.a.

The MRC-CMBI (http://www.imperial.ac.uk/cmmi) wishes to appoint two Lecturers/Senior Lecturers, equivalent in status to an Assistant Professor in the USA or a Group Leader in Europe, to direct a research programme in the area of bacterial pathogenesis and/or host-pathogen interactions. The successful candidates will have excellent opportunities to collaborate with the existing research groups within the MRC-CMBI and within Imperial College London.

The MRC Centre for Molecular Bacteriology and Infection (CMBI) is located at the South Kensington Campus, Imperial College London. The MRC-CMBI is a unique centre of excellence studying microbial pathogenesis with an emphasis on diseases caused by bacteria. The Centre is supported by MRC-funded facilities and a dedicated PhD training programme.

You must have a PhD (or equivalent) in a relevant field. You will have a proven track record in conducting high quality research, evidenced by high quality publications in peer-reviewed journals. You must have the ability to raise funding from a variety of sources and to direct a competitive independent research programme. You must also have the ability to communicate and inspire students as you will be expected to contribute to our undergraduate and postgraduate teaching programmes.

Our preferred method of application is online via our website http://www3.imperial.ac.uk/employment (please select "Job Search" then enter the job title or vacancy reference number, including spaces, **NS 2012 205 DPN**, into "Keywords"). Please complete and upload an application form as directed. Also please include your CV and a statement of research interests and future plans (up to two A4 pages).

Informal enquiries may be made to Professor Alain Filloux at a.filloux@imperial.ac.uk. Should you have any queries about the application process please contact Sandrine Nurboja at s.nurboja@imperial.ac.uk

POSITIVE ABOUT DISABLED PEOPLE

Closing date: 23 November 2012.

Stonewall
DIVERSITY CHAMPION

Committed to equality and valuing diversity. We are also an Athena Bronze SWAN Award winner, a Stonewall Diversity Champion and a Two Ticks Employer.

U229439R

Visit
www.naturejobs.com
to seriously improve your career prospects.

naturejobs
making science work

Quantitative Biology

UC San Diego

The University of California San Diego invites applications from outstanding candidates for multiple tenure-track or tenured faculty positions as part of a multi-year, campus-wide initiative to establish a preeminent program in the area of Quantitative Biology (qBio), emphasizing quantitative experimentation and theoretical analysis to study living systems. The qBio initiative builds on the interdisciplinary infrastructure provided by the NSF Center for Theoretical and Biological Physics, the NIH Center of Excellence in Systems Biology, the BioCircuits Institute, existing divisional strengths and various interdepartmental graduate programs. Candidates are encouraged to apply to each specific position that fits their interests; successful candidates may be affiliated with multiple academic units and will complement and participate in the development of quantitative biology across campus. All candidates must have earned a Ph.D. or equivalent degree and demonstrate a strong commitment to teaching at the undergraduate and graduate levels. The University is committed to excellence and diversity in its faculty and student body. Preference will be given to scholars with demonstrated excellence and creativity in research, scholarship, and commitment to diversity, equity and inclusion in higher education.

Cellular Dynamics and Pattern Formation (10-503): The Division of Physical Sciences is leading a broad-based search for a creative experimentalist for a tenure-track position at the Assistant Professor level, with focus on the dynamical aspects of cellular differentiation and multicellular development using novel physical approaches. Subject matters may range from organs and tissues in higher eukaryotes to biofilms and other microbial communities. Approaches might include advanced methods of in vivo imaging and novel functional probes to quantify molecular interactions and forces to dissect the spatiotemporal progression of morphological and physiological change, with the goal to elucidate the organizational principles of living systems from the sub-cellular to multi-cellular scales. While this search will be conducted University-wide, successful candidates will ultimately be appointed to a specific department within UC San Diego.

Quantitative Biological Networks (10-506): The Division of Biological Sciences is leading a broad-based search for a tenure-track faculty position in Quantitative Biological Networks at the Assistant Professor level. Candidates pursuing innovative research focused on exploring network interactions that control intra- and/or inter-cellular behavior, are encouraged to apply. Potential areas of interest include but are not limited to the use of cellular systems to explore spatial-temporal control principles underlying decision making in signaling and gene-regulatory networks, multicellular or organismal systems directed towards a quantitative understanding of how networks govern biological interactions within populations of cells or tissues, or in ecological communities that contain multiple species. Departmental affiliation will be tailored to the selected candidate.

Quantitative and Chemical Biology (10-445): The Department of Chemistry and Biochemistry in the Division of Physical Sciences seeks to hire a talented and creative tenure-track Chemical Biologist who combines quantitative analysis, synthesis, and biological systems. Interests include but are not limited to candidates who have a strong chemistry background and predictively probe or engineer living systems.

Soft Condensed Matter and Biological Physics (10-496): The Department of Physics in the Division of Physical Sciences seeks to hire a talented and creative experimentalist for a tenure-track Assistant Professor position in the area of Soft Condensed Matter and biological physics. The successful candidate is expected to use physical reasoning and quantitative tools to define the future of pure and applied biological research. The search is broad-based, and extends to individuals with purely scientific approaches as well as those that combine advanced instrument design with scientific discovery. All candidates must have a Ph.D. in Physics or a closely related field.

Systems/Quantitative Developmental Biology (10-491): The Section of Cell and Developmental Biology, division of Biological Sciences, invites applications for a faculty position in Developmental Biology at the tenure-track Assistant Professor or Associate Professor level. Candidates pursuing innovative research using systems, quantitative or dynamical approaches to investigate the generation, regeneration, or maintenance of a particular cell type, tissue, or organ are encouraged to apply.

Quantitative Biosystems Engineering (10-507): The Department of Bioengineering in the Jacobs School of Engineering seeks to hire a tenure-track systems bioengineer who combines engineering principles and methods, and modern biomedicine to develop quantitative models of living systems in normal and pathophysiology. Interests include but are not limited to candidates who have a strong engineering background and quantitatively probe or engineer living systems.

Salary will be commensurate with qualifications and based on UC pay scales. Review of applications will commence as early as November 1, 2012 and will continue until the positions are filled. Interested applicants must submit a cover letter, curriculum vitae with a list of publications, statement of research, statement of teaching, reprints of 3 to 5 representative publications, contact information for 3 to 5 references, and a separate statement that addresses past and/or potential contributions to and leadership in promoting diversity, equity and inclusion (see http://facultyequity.ucsd.edu/Faculty-Applicant-C2D-Info.asp). Applications must be submitted through the University of California San Diego's Academic Personnel RECRUIT System at https://apol-recruit.ucsd.edu/.

UCSD is an equal opportunity / affirmative action employer with a strong institutional commitment to excellence and diversity (http://diversity.ucsd.edu/).

NW229409R

nature COMMUNICATIONS

Assistant/Associate/Senior Editor (Physical Sciences), China

Nature Communications, Shanghai

Nature Publishing Group (NPG) is a division of Macmillan Publishers Ltd, dedicated to serving the academic, professional scientific and medical communities. For further information please visit www.nature.com

Nature Communications is an innovative multidisciplinary science journal, which launched in April 2010. The journal publishes high-quality research from the breadth of the natural sciences, and already has a strong track record of publications. The journal offers rapid publication in an online environment, with Open Access options.

We now have an exciting opportunity for an editorial position in the physical sciences based in our Shanghai office. The ideal candidate will have a PhD (or equivalent), and preferably postdoctoral experience, in chemistry and will be able to demonstrate a keen interest in the natural sciences in general. Editorial experience would be beneficial, but is not required, although applicants with significant editorial experience are encouraged to apply and will potentially be considered for Senior Editor positions.

Reporting to the Executive Editor, China (*Nature Communications*), the successful candidate will play an important role in determining the representation of their field in the journal, and will work closely with the other editors on all aspects of the editorial process, including manuscript selection and overseeing peer review. This role will also entail significant liaison with the scientific community, the Research Journal Chief Editors (UK and US) and the Editorial Advisory Panel. In addition, the successful candidate will be required to represent *Nature Communications* and *Nature Publishing Group* within the scientific community in China through lab visits and seminars and this will form a core part of the position.

This is a demanding and intellectually stimulating position. Broad scientific knowledge and training, excellent literary skills and a keen interest in the practice and communication of science are a prerequisite. The successful candidate must therefore have excellent communication and interpersonal skills. In addition the applicant should be fluent in English whilst skills in Mandarin would be advantageous (although not essential). The salary and benefits will be competitive, reflecting the critical importance and responsibilities of this role.

The role is offered on a full-time, permanent basis and is located in our Shanghai office following an expected 6 month training period in the London office (depending on previous experience).

Any offer of employment will be contingent on the necessary work permits for the UK and China, as applicable, being granted (the Company will provide reasonable assistance with the necessary application process in this respect).

Please note previous applicants should not apply.

Applicants should submit a CV, a covering letter explaining their interest in the post and a Research Highlights piece in the style of Nature Physics (300 words) on a recent paper from the literature to http://jobs.macmillan.com. Please also confirm your current salary and quote **NPG/157/11.**

Closing Date: 12th October 2012

nature publishing group npg

IN228974R

NYU Langone
MEDICAL CENTER

Chairperson, Pathology

The NYU School of Medicine has an opportunity for a Chair of the Department of Pathology to join the organization. The Dean and faculty consider this an exceptional opportunity to lead a preeminent institute in the City of New York in close collaboration with the other schools and colleges of New York University.

In this vital role, you will drive the vision, strategy and development of the department. Specifically, we will look to you to head the research and recruitment efforts. The successful candidate will have a PhD and/or MD degree, and will have demonstrated leadership experience in a university, medical center or research institute setting. In addition, you should have a distinguished record of scientific achievements, ability to foster collaborations and promote innovation and creativity. Expertise in the areas of immunology and/or genetics is preferred.

For additional information regarding the Department of Pathology at NYU please visit: **http://pathology.med.nyu.edu/**

Applications and nominations with accompanying curriculum vitae should be sent electronically, for confidential review by the search committee, to: **Andrea Botta, Project Manager for Education, Faculty and Academic Affairs, andrea.botta@nyumc.org**.

NYU Langone Medical Center is an equal employment/affirmative action employer and does not discriminate on the basis of race, color, religion, gender, gender identity or expression, sexual orientation, marital/or parental status, age, national origin, citizenship, disability, veterans status, or any other classification protected by applicable Federal, State, or Municipal Law.

NW229302R

Yale University
School of Medicine

FACULTY POSITION
AT THE JUNIOR OR SENIOR LEVEL

DEPARTMENT OF CELLULAR AND MOLECULAR PHYSIOLOGY

The Department of Cellular and Molecular Physiology is conducting a search for new faculty members at the junior level.

The search seeks candidates whose research connects the properties of molecules to the properties of physiological systems, with particular emphasis on integrative or translational research, or research that makes use of genetic techniques.

Excellent opportunities are available for collaborative research, as well as for graduate and medical student teaching. Candidates must hold a Ph.D., M.D., or equivalent degree. Applicants should include a curriculum vitae, a statement of research interests and goals, and should arrange to have three letters of reference sent. Applications should be emailed to **leisa.strohmaier@yale.edu** in PDF format.

Application Deadline: **November 16, 2012**

Yale University is an
Affirmative Action/Equal Opportunity Employer.

NW229101R

A joint medical school by Imperial College London and Nanyang Technological University, Singapore

LEE KONG CHIAN
SCHOOL OF
MEDICINE

NANYANG TECHNOLOGICAL UNIVERSITY | Imperial College London

Lee Kong Chian School of Medicine invites applications for:

- **Professor/Associate Professor, Integrative Physiology**
- **Professor/Associate Professor, Molecular Medicine**

Lee Kong Chian School of Medicine (LKCMedicine) is a new joint medical school founded by Imperial College London and Nanyang Technological University, Singapore. It was established in September 2010 and will operate from two campuses – Yunnan Campus at Nanyang Technological University and a new campus to be built at Novena. Its primary healthcare partner is the National Healthcare Group which comprises an integrated network of primary healthcare polyclinics, acute care hospitals such as Tan Tock Seng Hospital, and national specialty centres. Lee Kong Chian School of Medicine will deliver a new five-year undergraduate medical programme resulting in a joint degree (MBBS) awarded by Imperial College London and Nanyang Technological University. The first cohort of 50 students will start in August 2013. The School is building on the strengths of Imperial College London and Nanyang Technological University, Singapore to create a research-led centre of excellence for biomedical research and medical technologies.

The research strategy of the School builds on existing competencies at Nanyang Technological University and clinical opportunities at our healthcare partners, as well as areas of mutual interest between Imperial College London and Nanyang Technological University, both ranked within the top 50 global universities in the QS World rankings (6 and 47 respectively). The School's initial areas of research interest will be in Infectious Diseases including malaria, TB and microbiology, Metabolic Diseases including diabetes, Neuroscience and Mental Health, Medical and Surgical Technologies, and Health Services Research, and will incorporate the potential for synergistic interdisciplinary research with for example the School of Biological Sciences, The Singapore Centre on Environmental Life Sciences Engineering and the School of Chemical and Biomedical Engineering. We are planning two major capital developments both expected to open in 2015: 1) An Experimental Medicine Building at NTU, with first class research and education facilities, linked to the School of Biological Sciences, and 2) on the Novena campus, a Clinical Sciences Building containing state-of-the-art research and learning facilities immediately adjacent to Tan Tock Seng Hospital, an acute general hospital equipped with world-class medical facilities and an established reputation in the field of providing healthcare for the community. These facilities will provide an extraordinary new research environment.

LKCMedicine has an exciting new medical curriculum where the main learning experience will be based on Team-Based Learning in an integrated learning environment. Initial content for these courses is currently being developed jointly between Imperial College London and Nanyang Technological University.

The School now invites applicants with an international reputation in both education and research in Integrative Physiology / Molecular Medicine to apply for the positions, which will be based full-time in Singapore. This is an exciting opportunity to provide leadership in these areas in an innovative new medical school in a globally vibrant environment. Suitable candidates will be offered an Honorary Contract with Imperial College London.

The appointees will be the academic leads in the areas of Integrative Physiology and Molecular Medicine. The appointees will be expected to help to integrate the science relating to integrative physiology and molecular medicine in health and disease into the curriculum and to ensure the delivery of excellence in these areas in the Team-Based delivery of the curriculum. The appointees are also expected to undertake high quality internationally leading research in Integrative Physiology or Molecular Medicine in line with the research strategy of the School and in conjunction with other recent appointments, obtain research funding and supervise research staff and students. The appointees will be responsible for developing research activities, by sustaining a personal research plan, developing innovative research proposals and funding bids, managing and undertaking research activities, publishing in international peer reviewed journals, reviewing manuscripts and grant applications when appropriate.

If you are interested to pursue a career with the School, please apply to Ms Lydia Tan, Director, Human Resources (email: hr_LKCMedicine@ntu.edu.sg). For further queries, you may write to Professor Martyn Partridge, Senior Vice Dean of Lee Kong Chian School of Medicine (email: m.partridge@ntu.edu.sg / m.partridge@imperial.ac.uk).

Closing date: 15 November 2012

About Nanyang Technological University

A research-intensive public university, Nanyang Technological University (NTU) has 33,500 undergraduate and postgraduate students in the colleges of Engineering, Business, Science, and Humanities, Arts, & Social Sciences.

NTU is also home to four world-class autonomous institutes – the National Institute of Education, S. Rajaratnam School of International Studies, Earth Observatory of Singapore, and Singapore Centre on Environmental Life Sciences Engineering – and various leading research centres such as the Nanyang Environment & Water Research Institute (NEWRI) and Energy Research Institute @ NTU (ERI@N).

In recent years, Nanyang Technological University has established itself as a key player in Singapore's competitive research landscape, with a five-fold increase in its annualised amount of external competitive research grants won, from S$48M in FY2005 to S$255M in FY2010.

For more information, visit www.ntu.edu.sg

About Imperial College London

Consistently rated amongst the world's best universities, Imperial College London is a science-based institution with a reputation for excellence in teaching and research that attracts 14,000 students and 6,000 staff of the highest international quality. Innovative research at the College explores the interface between science, medicine, engineering and business, delivering practical solutions that improve quality of life and the environment - underpinned by a dynamic enterprise culture.

Since its foundation in 1907, Imperial's contributions to society have included the discovery of penicillin, the development of holography and the foundations of fibre optics. This commitment to the application of research for the benefit of all continues today, with current focuses including interdisciplinary collaborations to improve global health, tackle climate change, develop sustainable sources of energy and address security challenges.

For more information, visit www.imperial.ac.uk

JP229077R

UHH

Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

The **School of Mathematics, Informatics and Sciences** and the **Biocenter Klein Flottbek** invite applications for a

W1 JUNIOR PROFESSORSHIP IN MOLECULAR PLANT PHYSIOLOGY

commencing as soon as possible/**Ref.no. JP 188**.

Applicants are expected to demonstrate research experience on an international level as well as experience with externally funded projects. The University places particular value on the quality of teaching. We therefore request applicants provide details of both their teaching experience and their future teaching objectives.

The University aims to increase the number of women in research and teaching and explicitly encourages women to apply. Equally qualified female applicants will receive preference in accordance with Hamburg's Higher Education Act.

Tasks and responsibilities:
The successful candidate should have a strong background in molecular plant physiology and will be expected to research on cellular and/or sub-cellular regulation of physiological processes (for instance using imaging techniques). The post holder's research should strengthen the research focus of the Biocenter Klein Flottbek, the Department of Biology, and the School of Mathematics, Informatics and Sciences. Collaborations with the Center for Structural Systems Biology (CSSB) are conceivable. The successful candidate will be expected to integrate an externally funded research project into current and/or future joint research projects. The successful candidate will be expected to teach plant physiology on the School's bachelor's and master's degree programs in biology.

Requirements:
Academic qualifications and further requirements as specified in § 18 of Hamburg's Higher Education Act.

Severely disabled applicants will receive preference over equally qualified, non-disabled applicants.

For further information, please contact Prof. Dr. Stefan Hoth on +49 40 428 16-582 or at stefan.hoth@uni-hamburg.de.

Applications should include a curriculum vitae, copies of references and certificates, a list of publications, details of teaching experience and proposals for future courses and future research plans. The deadline for applications is **November 15th, 2012**. Please send applications, accompanied by the **reference number JP 188**, to Den Präsidenten der Universität Hamburg, Referat Organisation & Personalentwicklung, Ausschreibungsstelle, Mittelweg 177, 20148 Hamburg, Germany or via e-mail to: UniHHAusschreibungsstelle@verw.uni-hamburg.de.

Zertifikat seit 2010 audit familiengerechte hochschule

W229436R

Scientific & English Editor

INSTITUTE OF MOLECULAR BIOLOGY
ACADEMIA SINICA, TAIPEI, TAIWAN

The Institute of Molecular Biology at Academia Sinica, Taiwan's premier research institution, is currently recruiting a Scientific & English Editor. Responsibilities include substantive editing of manuscripts for publication in international scientific journals. Applicants should have a Ph.D. in Life Sciences (preferably molecular biology, cell biology or related areas), excellent literary skills, and be highly detail-oriented.

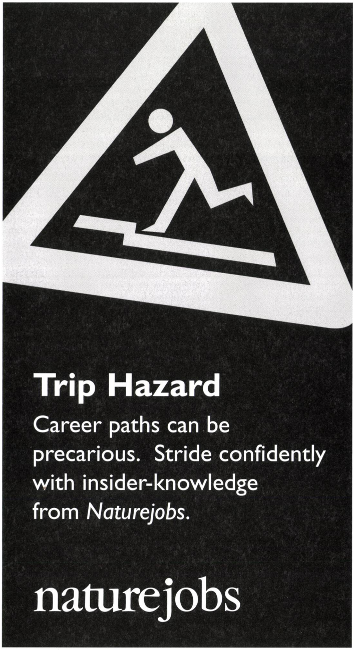
The selection process will start on October 15, 2012, and continue until the position is filled. We offer a competitive salary with benefits in a pleasant work environment. Applicants should send their resume and cover letter, stating their earliest availability and describing their suitability for this position to: Vivi Chiang, Institute of Molecular Biology, Academia Sinica, Nankang, Taipei, Taiwan 11529. E-mail: vivi@imb.sinica.edu.tw

JP229386R

"We recently placed an ad with naturejobs.com and the response was huge. We have now found the right candidate and will offer them the job"

Dr. Birgit Luber
Technische Universität
München, Germany

naturejobs


Trip Hazard
Career paths can be precarious. Stride confidently with insider-knowledge from *Naturejobs*.
naturejobs

Imperial College London

Division of Cell & Molecular Biology
Faculty of Natural Sciences

Postdoctoral Research Associate

£32,100 - £37,570 p.a.

Research Technician in Immunology

£28,200 - £32,100 p.a.

Positions as a Research Associate (Postdoctoral Fellowship) and a Research Technician in Immunology are available as soon as possible. This is an exciting opportunity to join a multidisciplinary team using novel and state-of-the-art microscopy for studying immune cell biology. The project involves using new fluorescence microscope technologies to study fundamental properties of T cell and/or Natural Killer cell immune synapses with a view to applying this knowledge to the design of drugs or therapeutic antibodies.

The successful candidates will be based in Professor Dan Davis's laboratory in Imperial College London, up to 31 March 2013. The group then relocates from 1 April 2013 to the University of Manchester; specifically the Manchester Collaborative Centre for Inflammation Research, a new Centre funded by the University of Manchester and pharmaceutical companies, where Professor Davis will be the Director of Research. You may be able to transfer before or after 1 April 2013, if desired.

Our preferred method of application is online via our website http://www3.imperial.ac.uk/employment Please select "Job Search" then enter the job title or **vacancy reference number including spaces – NS 2012 201 DPN (for the Research Associate post)** or **NS 2012 206 DPN (for the Research Technician post)** into "Keywords". Please complete and upload an application form as directed.

For informal enquiries please contact Professor Davis e-mail: d.davis@imperial.ac.uk

Alternatively, if you are unable to apply online, please contact Coralie Simmons on tel: 020 7594 2170 or e-mail: c.simmons@imperial.ac.uk to request an application form.

Closing date: 31 October 2012

Committed to equality and valuing diversity. We are also an Athena Bronze SWAN Award winner, a Stonewall Diversity Champion and a Two Ticks Employer.

POSITIVE ABOUT DISABLED PEOPLE

Stonewall
DIVERSITY CHAMPION

U229438R

You don't have to look this hard to find a job with Naturejobs.

naturejobs
making science work

IC CI
CNR Istituto di Cristallografia

CON IL SOSTEGNO DI
FONDAZIONE CON IL SUD

Three-year fixed-term contracts for young Italian researchers CNR - Istituto di Cristallografia, Bari-Italy

CNR-Istituto di Cristallografia, Bari announces the launch of an open call for two positions as Research Associate. The purpose of the call is to encourage young Italian researchers engaged in research activities abroad to return to Italy. Candidates must be of Italian nationality and with at least four years of research experience gained abroad in public or private institutions.

Prospective appointees are invited to send their applications for one of the following two positions:

The first research associate position will involve mainly the design of new chemical oligonucleotides and the devising of their synthesis. Experience is requested in the design and preparation of biological macromolecules such as natural and artificial PNA, DNA or chimeras, in asymmetric catalysis and in the synthesis and purification of organic multifunctional products.

The second research associate position will involve the development of a synthetic methodology to be applied to the preparation of designed compounds for medicinal chemistry studies. Experience is requested in multistep organic synthetic methodology, especially in the preparation of molecules in blocked conformation such as lactams, cyclopropanes or fused polycycles, in asymmetric catalysis, synthesis via multicomponent reactions and in the synthesis of macromolecular derivatives for drug delivery vehicles.

The research activity will be carried out within the project "Verso la medicina Personalizzata: Sviluppo di nuove molecole selettive per la cura del Neuroblastoma" approved by Fondazione CON IL SUD.

For more details, please read the full text of the announcement on the websites www.urp.cnr.it or www.ic.cnr.it.

W229465R

PANCREATIC CANCER UK
Striving for Survival

Applications now open for Pancreatic Cancer UK's Research Innovation Fund

We invite applications for innovative approaches to the biology, treatment and diagnosis of pancreatic cancer, including those successful in other areas of cancer that have justifiable promise for pancreatic cancer.

£½ million is available for the 1st round of the Fund in 2012. Successful awards will be funded for up to £75,000 for up to 12 months.

Closing date: 14th January 2013.

Visit **www.pancreaticcancer.org.uk/research** for more details about this and our **Future Leaders Fund** - clinical research training fellowships (PhD) jointly run with the Medical Research Council. Closing date: 15th January 2013.

Email queries to: **research@pancreaticcancer.org.uk**

Registered charity number: 1112708

U229287A

Infectious career articles each week

Career advice you can't put down.

naturejobs

REPROGRAPHICS: CC Media Group, London EC1V 8AR UK and **The Charlesworth Group,** Wakefield, UK.
PRINTED BY: William Gibbons & Sons Ltd, UK; **Publishers Press**, Lebanon Junction, Ky, USA and **Obun Printing Co. Inc,** Tokyo, Japan

Skolkovo Tech

Skolkovo Institute of Science and Technology

Call for International Research Center Proposals

Skolkovo Tech invites collaborative proposals from research and educational institutions around the world in the following areas:

- Biomedicine
- Energy
- Information Technology
- Nuclear
- Space

Skolkovo Tech is a private graduate research university located outside Moscow, Russia. Established in 2011 in collaboration with MIT and the Skolkovo Foundation, the university is founding 15 multidisciplinary Centers for Research, Education and Innovation (CREI) to address critical scientific and technology challenges.

With three centers in negotiations, the Second Round Call for Proposals offers an opportunity to partner with Russian investigators in a world-class research program.

Visit **www.skolkovotech.ru/research** for details.

SkTech/MIT
INITIATIVE

30 November 2012 deadline
$6-12M USD annually, up to five years

Scientific impact
Education
World-class research in Russia
Innovation

Skolkovo Tech is currently hiring tenure-track and tenured faculty. Details available at: **www.skolkovotech.ru/faculty**

a nature conference

2012 *Nature* / Institute for Genomic Medicine Symposium

Building upon the outstanding success of the Inaugural and 2011 Symposia at the Institute for Genomic Medicine at University of California, San Diego, this year's symposium, in partnership with *Nature*, is aimed at the international scientific community, with the goal of fostering and facilitating communication and collaborative initiatives in medical genomics.

The following topics will be covered at the conference: Cancer Genomics, Gene & Cell Therapies, Complex Human Traits, Personalized Genomics for Implementation of Genetic Information.

November 4-5, 2012
UCSD School of Medicine Campus
La Jolla, CA, USA

KEYNOTE SPEAKER:

Eric Green (National Human Genome Research Institute, NIH, USA)

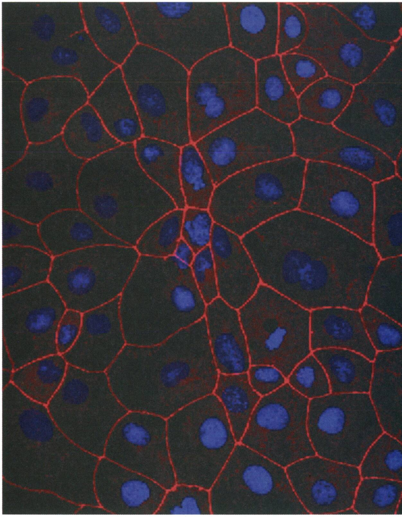
CONFIRMED SPEAKERS:

Robin Ali
Ines Barosso
Jean Bennett
George Church
Mark Daly
Richard Gilbertson
Joseph Gleeson
Joe Gray
David Haussler
Vanessa Hayes
Jonathan Sebat
Mike Snyder
Mike Stratton
Doug Turnbull

ORGANIZERS:

Kelly Frazer (University of California, San Diego, USA)
Xiang-dong Fu (University of California, San Diego, USA)
Magdalena Skipper (*Nature*, UK)
Kang Zhang (University of California, San Diego, USA)
Kun Zhang (University of California, San Diego, USA)

For more information and to register visit:
www.nature.com/natureconferences/nigm2012



IGM Institute for Genomic Medicine

UC San Diego
HEALTH SCIENCES

nature

TEMPORAL VENTURES ROBBED ME

Wish you were here?

BY SCOTT C. MIKULA

Posted: Nov. 17, 2029 8:03 PM EST

Note: Since TV DELETED my review on their website, I am reposting this on rateyourtrip.com where they can't touch it. It says something that they felt like they had to CENSOR me, don't you think?

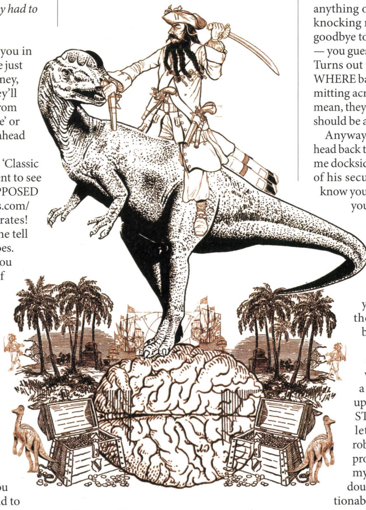
Don't let Temporal Ventures lure you in with their false advertising! They're just trying to cheat you out of your money, or if you're really lucky maybe they'll thump you on the head and pry it from your fingers in the name of 'science' or 'intertemporal law'. But I'm getting ahead of myself.

I enrolled in Temporal Venture's 'Classic Caribbean' package. Take a moment to see what the trip to the year 1720 is SUPPOSED to include: http://temporalventures.com/packages?p=39284. Treasure! Pirates! Adventure! Sound exciting? Let me tell you how the 'adventure' actually goes.

First off, the time jump makes you feel like SHIT, like after a night of drinking with my ex-wife. And when you finally crawl out of bed do they let you enjoy the sun? No, they have a two-hour orientation video. Maybe you'll make it past the first mention of temporal nodes or causality thresholds, but I didn't. Yawn. I didn't pay for a science lesson.

Then you meet wet-blanket Randy. He's the ~~tour guide~~ fun police for the trip. You know the type: little guy, glasses, with a stuck up, I'm-smarter-than-you attitude, always monotoning about what you should or shouldn't be doing. I had to spend most of my morning tuning him out. When I asked about treasure hunting, stick-up-his-butt Randy made some excuse about liability and pointed me to the 'pirate's wreck' tour — some old ship that got caught up on a reef and I guarantee you doesn't have an ounce of treasure on it by now. Meanwhile all the other guests are content to lay on the beach or sit at the bar and blab about how clear the sky is and that the air smells so fresh. Whatever. Come home with a load of Spanish treasure and you can invest in a personal holodome and next-gen atmo-filter. You'll be able to lay on a beach whenever you want.

➲ NATURE.COM
Follow Futures on Facebook at:
go.nature.com/mtoodm



JACEY

Oh, and pirates? There was ONE pirate. Greasy Sid had figured out how to panhandle from the tourists — which goes to show that bums will be bums no matter what century it is — and I guess some intertemporal treaty says they're not legally able to run him off. I think he had tuberculosis or something, but I talked to him anyway. Well, with the right incentive (hint: pirates like Bacardi Gold way better than whatever booze they're used to) HE, at least, was willing to talk about treasure. Even drew me a map on the back of a cocktail napkin.

So I, er, 'borrowed' some scuba gear and a skimmer. And damned if Greasy Sid didn't know what he was talking about. I had to swim through an underwater cave and almost got skewered by poisoned darts and I probably set off some ancient voodoo curse, but I found it. Treasure like you've never seen! Chests and sacks of coins, heaps of gems, cups and plates of gold and silver, and enough jewellery to make my ex-wife's eyes pop.

I tried to send my ex a picture with a little note about how she thought I'd never make anything of myself and as she was always knocking my 'schemes', she could just say goodbye to seeing any part of the haul, but — you guessed it, my cell showed no signal. Turns out they don't have coverage ANYWHERE back then. Something about transmitting across the time link, I don't know. I mean, they're scientists, for god's sake! They should be able to figure something out.

Anyway, I bundle up what I can carry, head back to the resort, and guess who meets me dockside? Sonovabitch Randy, with two of his security goons flanking him. "You know you can't take any of that home with you," he says, all smug, and then he blabs on about disrupting the time flow and I point out that he told us a dozen times that nothing we do here can mess up the future, and he says: "That's true, as long as you don't take anything FROM the past TO the future. Then all bets are off."

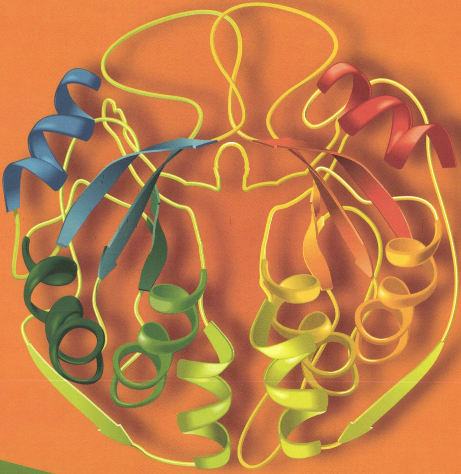
Which you'd THINK someone would have mentioned if it was such a big deal. But it's just a bunch of bullshit that he made up so he could 'confiscate' *cough* STEAL *cough* my treasure. So let's review. Temporal Ventures robbed me of my rightfully obtained property, manhandled me, searched my person — they even stole the doubloon I'd stuffed into my unmentionables because I KNEW they'd pull some crap like that — and then they had security glued to me for the rest of the trip so I couldn't even take a piss without one of them watching over my shoulder.

Well, if THAT'S what you want from your 'vacation' then go right ahead with Temporal Ventures. They've lost my business. I was looking forward to checking out their 'Jurassic Jaunt' package, but screw that. I bet they don't even let you ride the dinosaurs. ■

Scott C. Mikula *is a husband, father, software developer, board gamer, and has been told he makes a mean chocolate-chip cookie. On his better days he is also a writer.*

7,000 Human Proteins Made from Human Cells.

More Assays! Better Assays!



The Largest Collection of Mammalian Cell-Derived Proteins

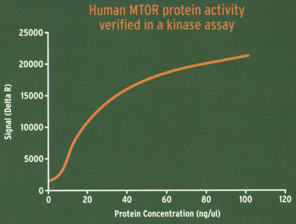
Protein functional studies

Compound screening

Immunogens for antibody production

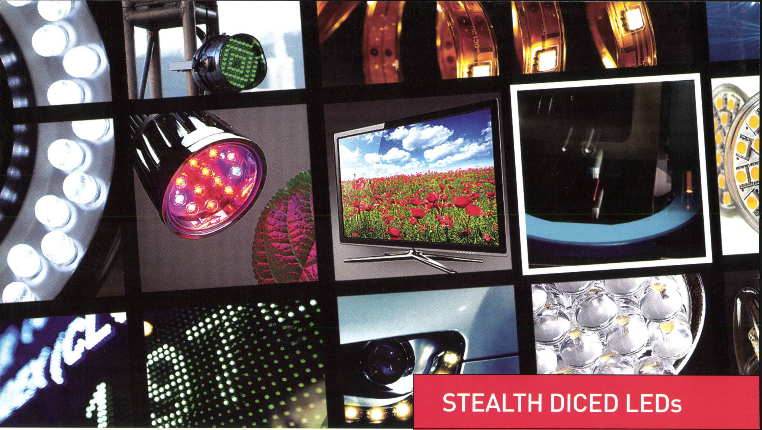
Assay standards: ELISA, Mass Spec, and more

Bulk production available



ORIGENE
Your Gene Company

origene.com/protein



STEALTH DICED LEDs

BRIGHTER WORLD

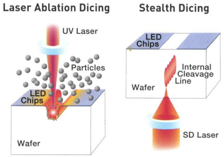
Hamamatsu Photonics "Stealth Dicing" technology helps produce better, brighter, more affordable LEDs to enlighten our world.

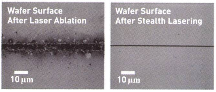
Cutting an LED semiconductor wafer into individual chips with a laser, as manufacturers generally do, produces heat and destroys some of the reflective surface layers that help to channel light. This results in dimmer LEDs.

Hamamatsu Photonics solved these problems by reinventing a decade-old technology named Stealth Dicing. With Stealth Dicing, instead of cutting across the surface of the wafer, the laser is focused beneath the surface, creating a precise, "hidden" rift. The chips then break apart neatly.

No water is needed, saving a precious resource. Also, the width of the rift is nearly zero, sparing the reflective surface and thus making brighter LEDs for less money.

These improved efficiencies open the door for widespread use of these energy saving lights in general illumination and paves the way for other applications of Stealth Dicing. For example, stealth-diced silicon carbide chips could be used in next generation electronic control systems in electric cars, photovoltaics and smart grids.


Laser Ablation Dicing
UV Laser
Particles
LED Chips
Wafer
Stealth Dicing
LED Chips
Internal Cleavage Line
Wafer
SD Laser


Wafer Surface After Laser Ablation
10 µm
Wafer Surface After Stealth Lasering
10 µm

Stealth Dicing eliminates the molten debris of laser ablation dicing, so it requires no water clean-up.

Hamamatsu Photonics' expertise with light has revealed an elegant solution to a complex problem.

HAMAMATSU

PHOTON IS OUR BUSINESS

For more information, visit our site at: http://jp.hamamatsu.com/en/rd/publication

HAMAMATSU PHOTONICS K. K.

325-6, Sunayama-cho, Naka-ku, Hamamatsu City, 430-8587 Japan Telephone: (81)53-452-2141, Fax: (81)53-456-7889 URL http://www.hamamatsu.com